General Internal Medicine
Boston University School of Medicine
2008 Publications - A-K

1. Aggarwal A, **Freund K**, Sato A, Adams-Campbell LL, Lopez AM, Lessin LS, Ockene J, Wallace RB, Williams CD, Bonds DE. 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 (Larchmt). 2008;17(8):1353-61.
2. Albers AB, Biener L, Siegel M, **Cheng DM**, Rigotti N. Household smoking bans and adolescent antismoking attitudes and smoking initiation: Findings from a longitudinal study of a massachusetts youth cohort. Am J Public Health. 2008;98(10):1886-93.
3. **Alford DP, Liebschutz J**, Chen IA, Nicolaidis C, Panda M, Berg KM, Gibson J, Picchioni M, Bair MJ. Update in pain medicine. J Gen Intern Med. 2008;23(6):841-5.
4. **Alford DP**, Richardson JM, **Chapman SE**, Dube CE, Schadt RW, **Saitz R**. A web-based alcohol clinical training (ACT) curriculum: Is in-person faculty development necessary to affect teaching? BMC Med Educ. 2008;8:11.
5. Ankle Brachial Index Collaboration: 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, Rodriguez 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(2):197-208.
6. Apter AJ, **Paasche-Orlow MK**, Remillard JT, Bennett IM, Ben-Joseph EP, Batista RM, Hyde J, Rudd RE. Numeracy and communication with patients: They are counting on us. J Gen Intern Med. 2008;23(12):2117-24.
7. Bussey HI, Wittkowsky AK, **Hylek EM**, Walker MB. Genetic testing for warfarin dosing? not yet ready for prime time. Pharmacotherapy. 2008;28(2):141-43.
8. **Carr P**, Woodson J. Academic performance of ethnic minorities in medical school. BMJ. 2008;337:a1094.
9. Chonchol M, Whittle J, Desbien A, Orner MB, Petersen LA, **Kressin NR**. Chronic kidney disease is associated with angiographic coronary artery disease. Am J Nephrol. 2008;28(2):354-60.
10. Clark NP, Witt DM, Delate T, Trapp M, Garcia D, Ageno W, **Hylek EM**, Crowther MA, 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):960-67.
11. Clay BJ, Halasyamani L, Stucky ER, **Greenwald JL**, Williams MV. Results of a medication reconciliation survey from the 2006 society of hospital medicine national meeting. J Hosp Med. 2008;3(6):465-72.
12. Craft LL, Perna FA, **Freund KM**, Culpepper L. Psychosocial correlates of exercise in women with self-reported depressive symptoms. J Phys Act Health. 2008;5(3):469-80.

13. Edwards EM, **Cheng DM**, Levenson S, **Bridden C**, **Meli S**, Egorova VY, Krupitsky EM, **Samet JH**. Behavioral assessments in Russian addiction treatment inpatients: A comparison of audio computer-assisted self-interviewing and interviewer-administered questionnaires. HIV Clin Trials. 2008;9(4):247-53.

14. Fox CS, Pencina MJ, D'Agostino RB, **Murabito JM**, Seely EW, Pearce EN, Vasan RS. Relations of thyroid function to body weight: Cross-sectional and longitudinal observations in a community-based sample. Arch Intern Med. 2008;168(6):587-92.

15. **Freund KM**, **Battaglia TA**, Calhoun E, Dudley DJ, Fiscella K, Paskett E, Raich PC, Roetzheim RG, Patient Navigation Research Program Group. National cancer institute patient navigation research program: Methods, protocol, and measures. Cancer. 2008;113(12):3391-99.

16. Friedman BW, **Kapoor A**, Friedman MS, Hochberg ML, Rowe BH. The Relative Efficacy of Meperidine for the Treatment of Acute Migraine: A Meta-analysis of Randomized Controlled Trials. American College of Emergency Physicians. 2008;52(6):705-13. doi:10.1016.

17. Garcia DA, Regan S, **Henault LE**, Upadhyay A, Baker J, Othman M, **Hylek EM**. Risk of thromboembolism with short-term interruption of Warfarin therapy. Arch Intern Med. 2008;168(1):63-69.

18. Garcia D, Witt DM, **Hylek EM**, 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 July/Aug 2008;42:979-88. doi:10.1345/aph.1L098.

19. Gordon AJ, Fiellin DA, Friedmann PD, Gourevitch MN, Kraemer KL, Arnsten JH, **Saitz R**, Society of General Internal Medicine's Substance Abuse Interest Group. Update in addiction medicine for the primary care clinician. J Gen Intern Med. 2008;23(12):2112-16.

20. Govindaraju DR, Cupples LA, Kannel WB, O'Donnell CJ, Atwood LD, D'Agostino RB S, Fox CS, Larson M, Levy D, **Murabito J**, Vasan RS, Splansky GL, Wolf PA, Benjamin EJ. Genetics of the framingham heart study population. Adv Genet. 2008;62:33-65.

21. **Hanchate AD**, **Ash AS**, Gazmararian JA, Wolf MS, **Paasche-Orlow MK**. The demographic assessment for health literacy (DAHL): A new tool for estimating associations between health literacy and outcomes in national surveys. J Gen Intern Med. 2008;23(10):1561-66.

22. **Hanchate AD**, Zhang Y, Felson DT, **Ash AS**. Exploring the determinants of racial and ethnic disparities in total knee arthroplasty: Health insurance, income, and assets. Med Care. 2008;46(5):481-88.

23. Hart RG, Pearce LA, Halperin JL, **Hylek EM**, Albers GW, Anderson DC, Connolly SJ, Friday GH, Gage BF, Go AS, Goldstein LB, Gronseth G, Lip GYH, Sherman DG, Singer DE, van Walraven C. Comparison of 12 risk stratification schemes to predict stroke in patients with non-valvular atrial fibrillation. Stroke. 2008;39:1901-10.

24. Hironaka LK, **Paasche-Orlow MK**. The implications of health literacy on patient-provider communication. Arch Dis Child. 2008;93(5):428-432.

25. **Hylek EM**. Oral anticoagulant therapy for patients with atrial fibrillation--an update. Thromb Res. 2008;123(S1):46-9.

26. **Hylek EM**. Contra: 'warfarin should be the drug of choice for thromboprophylaxis in elderly patients with atrial fibrillation'. caveats regarding use of oral anticoagulant therapy among elderly patients with atrial fibrillation. Thromb Haemost. 2008;100(1):16-17.

27. **Hylek EM**, Becker RC. Atrial fibrillation and stroke severity: Expanding the mechanistic exemplar, clinical phenotype, and goals of anticoagulant pharmacotherapy. Ann Neurol. 2008;64(5):480-83.

28. **Hylek EM**, Frison L, **Henault LE**, Cupples A. Disparate stroke rates on warfarin among contemporaneous cohorts with atrial fibrillation: Potential insights into risk from a comparative analysis of SPORTIF III versus SPORTIF V. Stroke. 2008;39(11):3009-14.

29. **Hylek EM**, Magnani JW. Methodological considerations for interpretation of rates of major haemorrhage in studies of anticoagulant therapy for atrial fibrillation. Europace. 2008;10(1):3-5.

30. **Hylek EM**, Solarz DE. Dual antiplatelet and oral anticoagulant therapy: Increasing use and precautions for a hazardous combination. JACC Cardiovasc Interv. 2008;1(1):62-64.

31. Jagsi R, Tarbell NJ, **Henault LE**, Chang Y, **Hylek EM**. The representation of women on the editorial boards of major medical journals: A 35-year perspective. Arch Intern Med. 2008;168(5):544-48.

32. Joseph NP, Hunkali KB, Wilson B, Morgan E, Cross M, **Freund KM**. Pre-pregnancy body mass index among pregnant adolescents: Gestational weight gain and long-term post partum weight retention. J Pediatr Adolesc Gynecol. 2008;21(4):195-200.

33. Kannel WB, Evans JC, Piper S, **Murabito JM**. Angina pectoris is a stronger indicator of diffuse vascular atherosclerosis than intermittent claudication: Framingham study. J Clin Epidemiol. 2008;61(9):951-57.

34. **Kapoor A**, Kader B, Cabral H, **Ash AS**, **Berlowitz D**. Using the case mix of pressure ulcer healing to evaluate nursing home performance. Am J Med Qual. 2008;23(5):342-49.

35. **Kartha A**, Brower V, **Saitz R**, **Samet JH**, Keane TM, **Liebschutz J**. The impact of trauma exposure and post-traumatic stress disorder on healthcare utilization among primary care patients. Med Care. 2008;46(4):388-93.

36. Katz RV, Claudio C, **Kressin NR**, Green BL, Wang MQ, Russell SL. Willingness to participate in cancer screenings: Blacks vs Whites vs Puerto Rican Hispanics. Cancer Control. 2008;15(4):334-43.

37. Katz RV, Green BL, **Kressin NR**, Kegeles SS, Wang MQ, James SA, Russell SL, Claudio C, McCallum JM. The legacy of the Tuskegee syphilis study: Assessing its impact on willingness to participate in biomedical studies. J Health Care Poor Underserved. 2008;19(4):1168-80.

38. Katz RV, Kegeles SS, **Kressin NR**, Green BL, James SA, Wang MQ, Russell SL, Claudio C. Awareness of the Tuskegee syphilis study and the US presidential apology and their influence on minority participation in biomedical research. Am J Public Health. 2008;98(6):1137-42.

39. Katz RV, Wang MQ, Green BL, **Kressin NR**, Claudio C, Russell SL, Sommervil C. Participation in biomedical research studies and cancer screenings: Perceptions of risks to minorities compared with whites. Cancer Control. 2008;15(4):344-51.

40. Kernan WN, **Hershman W**, Alper EJ, Lee MY, Viscoli CM, Perry JR, O'Connor PG. Disagreement between students and preceptors regarding the value of teaching behaviors for ambulatory care settings. Teach Learn Med. 2008;20(2):143-50.

41. **Kressin NR**, Jones JA, Orner MB, Spiro A,3rd. A new brief measure of oral quality of life. Prev Chronic Dis. 2008;5(2):A43.

42. **Kressin NR**, Raymond KL, **Manze M**. Perceptions of race/ethnicity-based discrimination: A review of measures and evaluation of their usefulness for the health care setting. J Health Care Poor Underserved. 2008;19(3):697-730.

43. **Kronman AC**, **Ash AS**, **Freund KM**, **Hanchate A**, Emanuel EJ. Can primary care visits reduce hospital utilization among Medicare beneficiaries at the end of life? J Gen Intern Med. 2008;23(9):1330-35.

44. Kuo TC, Zhao Y, Weir S, Kramer MS, **Ash AS**. Implications of co-morbidity on costs for patients with Alzheimer disease. Med Care. 2008;46(8):839-46.

JOURNAL OF WOMEN'S HEALTH
Volume 17, Number 8, 2008

DOI: 10.1089/jwh.2007.0544

Are Depressive Symptoms Associated with Cancer Screening and Cancer Stage at Diagnosis among Postmenopausal Women? The Women's Health Initiative Observational Cohort

Arpita Aggarwal, M.D., M.Sc.,[1] Karen Freund, M.D., M.P.H.,[2] Alicia Sato, M.S.,[3]
Lucille L. Adams-Campbell, Ph.D.,[4] Ana Maria Lopez, M.D., M.P.H., FACP,[5]
Lawrence S. Lessin, M.D., MACP,[6] Judith Ockene, Ph.D., M.Ed.,[7] Robert B. Wallace, M.D., M.Sc.,[8]
Carla D. Williams, Ph.D.,[4] and Denise E. Bonds, M.D., M.P.H.[9]

Abstract

Background: Women with depressive symptoms may use preventive services less frequently and experience poorer health outcomes. We investigated the association of depressive symptoms with breast and colorectal cancer screening rates and stage of cancer among a cohort of postmenopausal women.
Methods: In The Women's Health Initiative Observational Study, 93,676 women were followed on average for 7.6 years. Depressive symptoms were measured at baseline and at 3 years using the 6-item scale from the Center for Epidemiological Studies Depression scale (CES-D). We calculated a cancer screening rate expressed as a proportion of the years that women were current with recommended cancer screening over the number of follow-up visits in the study. Breast and colorectal cancers were staged based on Surveillance, Epidemiology and End Results (SEER) classification.
Results: At baseline, 15.8% (12,621) women were positive for depressive symptoms, and 6.9% (4,777) were positive at both baseline screening and at 3 years. The overall average screening rate was 71% for breast cancer and 53% for colorectal cancer. The breast cancer screening rate was 1.5% (CI 0.9%–2.0%) lower among women who reported depressive symptoms at baseline than among those who did not. Depressive symptoms were not a predictor for colorectal cancer screening. Stage of breast and colorectal cancer was not found to be associated with depressive symptoms after adjusting for covariates.
Conclusions: Among a healthy and self-motivated cohort of women, self-reported depressive symptoms were associated with lower rates of screening mammography but not with colorectal cancer screening.

Introduction

BREAST AND COLORECTAL CANCER are the second and third leading causes of cancer death among women in the United States.[1] Early detection of these cancers can save lives, reduce length of treatment, and increase quality of life (QOL). Race, economic status, family history of cancer, medical comorbidity, healthcare access, health behavior, and education have been recognized as major factors associated with screening behavior.[2,3] Only a few studies have investigated the role of psychiatric comorbidities in cancer screening behavior and mortality.

[1]Virginia Commonwealth University, Richmond, Virginia.
[2]Boston University, Boston, Massachusetts.
[3]Fred Hutchinson Cancer Research Center, Seattle, Washington.
[4]Howard University Cancer Center, Washington, DC.
[5]Arizona Cancer Center, University of Arizona, Tucson, Arizona.
[6]Washington Hospital Center, Washington, DC.
[7]University of Massachusetts Medical School, Worcester, Massachusetts.
[8]University of Iowa, Iowa City, Iowa.
[9]University of Virginia, Charlottesville, Virginia.
Preliminary results of this study were published in abstract format and subsequently presented as a poster at the annual meeting of the Society for General Internal Medicine in April 2006.
The WHI study was supported by the National Heart, Lung and Blood Institute and the General Clinical Research Center program of the National Center for Research Resources, Department of Health and Human Services.

In the United States, about 1 in 8 women can expect to develop clinical depression during her lifetime,[4,5] a condition that may cause considerable impairment, suffering, and disruption of personal, family, and work in one's life. Although depressed women are more likely to experience functional impairment,[5,6] less than half seek medical care.[7] Several studies have addressed the association of chronic mental illnesses, such as depression, with use of general medical care.[8–12] These studies indicate poor adherence to medical treatment and follow-up as well as worse outcome in the population with psychiatric illness or substance use disorder or dual diagnoses. Few of these studies looked specifically at preventive care and cancer screening among patients with chronic mental illnesses.[8,13] Although each of these studies shows lower cancer screening rates among patients with chronic mental illnesses, they did not study depression as a factor independent of substance abuse. The inclusion of individuals with dual diagnoses does not allow us to understand the independent effect of the mental illness.

Moreover, these studies included a broad spectrum of psychiatric illnesses, such as depression and schizophrenia, that are significantly different in their prevalence, latency, course of disease, and, most importantly, their effect on patient functioning. It may not be reasonable to study them as a single group of diseases. Furthermore, most studies were performed on veterans,[8–10] who exhibit higher rates of multiple mental illnesses and may experience a more serious course of disease compared with nonveterans.[14–16] There are no data available on the association of depression with cancer stage at initial presentation. We hypothesize that depressive symptoms may be associated with lower cancer screening rates as a result of reduced motivation for use of preventive services, less receipt of recommendations for screening by providers who are focusing on depression treatment, or reduced compliance with screening recommendations. These lower rates of early cancer screening may result in a more advanced stage of cancer at the initial presentation. On the other hand, depression itself might affect on the growth of cancer cells, leading to higher incidence of advanced stage cancers. Advanced stage of breast and colorectal cancers is associated with increased cancer morbidity, mortality, and healthcare costs. To better understand the association of depressive symptoms with cancer screening rates and stage of cancer at clinical presentation, we analyzed data from the Women's Health Initiative Observational Study (WHI-OS).

Materials and Methods

Study population

The WHI-OS is a cohort of 93,676 women with an average follow-up time of 7 years. Postmenopausal women aged 50–79 years who gave written informed consent were recruited into the WHI study at 40 clinical centers in the United States, mostly through mass mailings to age-eligible women. The WHI cohort is multiethnic, with 83.3% white, 8.2% African American, 3.9% Hispanic, 2.9% Asian/Pacific Islander, 0.5% American Indian/Alaskan Native, and 1.4% unknown ethnicity. Details of the WHI study design are reported elsewhere.[17] Exclusions for WHI-OS were participation in other randomized trials, survival prediction of <3 years, alcohol abuse, dementia, drug dependency, documented diagnosis of a serious mental illness (which includes schizophrenia, schizoaffective disorder, bipolar-affective disorder, and other nonorganic psychotic disorders), or other conditions making women unable to participate in the study. Because individuals with a prior history of cancer often have increased surveillance and sometimes have different screening guidelines, this analysis excluded all subjects with a history of cancer at baseline, except nonmelanoma skin cancer. Previous research has shown an increased incidence of depression among patients diagnosed with cancer.[18] Women with cancer diagnosed within the first year of the study were also excluded to reduce the bias of a causal association of cancer and depression.

Variables

Participants were assessed for their current and past history of depressive symptoms using the Burnam screen[19] for depression. This screen consists of 6 items from the 20-item Center for Epidemiological Studies Depression scale (CES-D) and 2 items from the National Institute of Mental Health's Diagnostic Interview Schedule (DIS).[20]

The 6-item CES-D and DIS scale was administered at baseline and again at the 3-year follow-up visit. Current depressive symptoms were assessed using 6 items from the CES-D in the Burnam scale, which is highly correlated with the 20-item CES-D scale (correlation coefficient $r = 0.88$, $p < 0.001$).[20,21] Burnam et al.[19] showed that this screen has adequate psychometric properties for detecting current depressive disorder (major depression and dysthymia), with 86% sensitivity and 95% specificity for detecting depression in a primary care population. Participants are asked how often they felt the depressive symptoms during the past week. Each item is scored as 0 (rarely or none of the time <1 day), 1 (some or a little of the time 1–2 days), 2 (occasionally or a moderate amount of the time 3–4 days), or 3 (most or all of the time 5–7 days). The items included: (1) you felt depressed, (2) your sleep was restless, (3) you enjoyed life (reversed scoring), (4) you had crying spells, (5) you felt sad, and (6) you felt that people disliked you.

The DIS scale consists of two questions used to assess depressive symptoms[20,21] over the previous 2 years. These dichotomous response questions are (1) In the past year, have you had 2 weeks or more during which you felt sad, blue, or depressed or lost pleasure in things that you usually cared about or enjoyed? (2) Have you had 2 years or more in your life when you felt depressed or sad most days, even if you felt okay sometimes? If yes, have you felt depressed or sad much of the time in the past year?"

A score of ≥ 5 on the short form of the CES-D at baseline was used as the primary definition for depressive symptoms. This definition represents depressive symptoms over the previous week at study baseline. During this longitudinal study, depressive symptoms were also assessed using another measure (DIS scale) and at a different time point (3 years). We used additional definitions for depressive symptoms to assess the correlation of depressive symptoms with cancer screening and stage at initial presentation. The three additional definitions in this sensitivity analysis used to strengthen our results are (1) depressive symptoms defined as both a score of ≥ 5 on the CES-D and a positive DIS score (score of 2) at baseline, (2) a score of ≥ 2 on the DIS, and (3) depressive symptoms defined as a score of ≥ 5 at both the baseline visit and 3 years later on the CES-D.

Current breast cancer screening was defined as a mammogram within the last 12 months.[22,23] Current colorectal screening was defined as an annual fecal occult blood test (FOBT) or lower endoscopy or double-contrast barium enema within the last 5 years.[22,23] This information was collected during the baseline and annual follow-up questionnaires. We calculated a screening rate expressed as a proportion of the years that women were current with recommended screening over the number of follow-up visits in the study.

Breast and colorectal cancers were staged based on the Surveillance Epidemiology and End Results (SEER) classification.[24] All *in situ* and localized cancers were classified as early stage cancers, and regional and distant cancers were classified as late stage cancers. Unstaged cancers and women diagnosed with cancer during the first year of the study were not included in this analysis, but there was little difference in the incidence rates for this subset of cancers between depressed and nondepressed cohorts (data not shown).

Sociodemographic characteristics, past medical history, and information about known breast and colorectal cancer risk factors were self-reported on the baseline questionnaire. Descriptive characteristics included age, ethnicity, education, income, insurance type, physical activity, age at menarche, age at first pregnancy, number of children breastfed, family history of breast or colorectal cancer, history of breast biopsy, history of ulcerative colitis or Crohn's disease, use of hormone therapy (HT) (estrogen only or estrogen and progesterone combination therapy), aspirin use, smoking and alcohol use status, and use of a primary care provider. Body mass index (BMI) was calculated based on weight and height measurements taken by study nurses at baseline. New medical problems and changes in treatment were reported during follow-up questionnaires. Comorbidity burden was calculated using a modified version of the Charlson Index (unpublished WHI data by R. Gold et al.), a commonly used and validated[25,26] comorbidity index composed of 19 diseases weighted based on how well they predict mortality, with a maximum possible score of 37. Charlson Index scores were calculated using WHI baseline data from each study subject's self-reported medical history. Use of antidepressant medication was not included in the model because of the variety of other indications for these classes of medication.

Statistical analysis

Baseline sociodemographic characteristics and cancer risk factors of the study cohort were compared among women who screened positive for depression vs. those who did not. Simple linear regression was used to determine the association between depressive symptoms and subsequent breast or colorectal cancer screening rates. To control for potential confounding factors, the estimates and corresponding 95% confidence intervals (95% CI) for depression status in the breast cancer screening model were adjusted for sociodemographic characteristics, family history of breast cancer, history of previous breast biopsy, HT, alcohol intake, BMI, comorbidity index, insurance, and having a primary care provider. Adjustment factors for the colorectal cancer screening analysis included sociodemographic characteristics; family history of colorectal cancer, ulcerative colitis, or Crohn's disease; alcohol intake; comorbidity index; and having a primary care provider.

Logistic regression was used to assess the correlation of depressive symptoms with late vs. early stage at presentation of subsequent breast and colorectal cancers. To control for potential confounding, odds ratios (ORs) and 95% CIs for depression status in the breast cancer model were adjusted for sociodemographic characteristics, insurance type, breast biopsy, number of relatives with breast cancer, moderate or strenuous physical activity, BMI, age at first birth, number of children breastfed, parity, and aspirin use. Depression status in the colorectal cancer model was adjusted for sociodemographic characteristics, insurance type, BMI, moderate or strenuous physical activity, and smoking status. We ran additional multivariate models with different measurements of depressive symptoms at baseline and 3-year follow-up. All analyses were conducted using SAS (version 9.1.3) (SAS Institute, Cary, NC) software. Analyses were statistically significant at alpha of 0.05.

Results

There were 12,621 (15.8%) women (Table 1) with current depressive symptoms and a mean CES-D score of ≥5 at baseline. The mean CES-D score for those above the cutoff of 5 was 7.0 (standard deviation [SD] 2.4), compared with 1.45 (SD 1.33) for women scoring below the cutoff. Using our alternative definitions, 5,152 (7.4%) women had both positive CES-D and DIS scores at baseline, 9,760 (12.1%) had positive DIS scores, and 4,777 (6.9%) women had positive CES-D at baseline and at the 3-year follow-up.

Table 2 compares the sociodemographic and health characteristics of women with and without depressive symptoms at baseline in the WHI-OS cohort. Women with depressive symptoms were younger (aged 50–59 years); had lower educational attainment; and were less likely to be white, have a >$20,000 annual income, and be insured. Women with depressive symptoms were more likely to have BMI ≥30, were less physically active, used more alcohol in the past, and were more likely to be current smokers.

The overall average screening rate was 71% for breast can-

Table 1. Prevalence of Depressive Symptoms Reported by Women: Women's Health Initiative Observational Cohort

Depressive symptoms	n	%	*Mean depressive score*	*SD*[a]
CES-D > 5 at baseline	12,621	15.8	7.0	2.4
DIS > 2 at baseline	9,760	12.1	8.0	2.5
CES-D > 5 and DIS > 2 at baseline	5,152	7.4	8.0	2.7
CES-D > 5 at baseline and 3-year follow-up	4,777	6.9	7.5	2.6

[a]CES-D, Center for Epidemiological Studies Depression scale; DIS: Diagnostic Interview Schedule; SD, standard deviation.

Table 2. Baseline Characteristics of Women with and without Depressive Symptoms: Women's Health Initiative Observational Cohort[a]

	Depressive symptoms[b] n = *12,623*		*No depressive symptoms* n = *67,368*	
Variable	n	%	n	%
Age group, at screening, years				
50–59	4,755	38	21,459	32
60–69	5,201	41	30,058	45
70–792,667	21	15,851	23	
Education				
High school diploma or less	3,402	27	13,413	20
Posthigh school/some college	4,836	39	24,109	36
College degree or more	4,262	34	29,324	43
Annual household income (dollars/year)				
<$20,000	2,794	24	8,660	14
$20,000–$50,000	5,088	44	27,061	43
$50,000+	3,736	32	26,830	43
Ethnicity				
White	9,998	79	56,730	84
Black	1,223	10	5,181	8
Hispanic	854	7	2,134	3
American Indian	80	<1	263	<1
Asian/Pacific Islander	262	2	2,147	3
Unknown	206	2	913	1
Body mass index (kg/m^2)				
<25	4,356	35	28,218	42
25–<30	4,219	34	22,738	34
≥30	3,900	31	15,631	24
Insurance				
None	747	6	1,951	3
Private only	6,488	52	36,145	54
Medicare/Medicaid only	1,189	10	4,822	7
Other	4,021	32	23,815	36
Physical activity				
No activity	2,425	19	8,273	12
Some activity	5,217	42	25,009	38
2–<4 episodes per week	2,049	16	12,632	19
4+ episodes per week	2,839	23	20,837	31
Alcohol intake				
Never drinker	1,383	12	7,516	11
Past drinker	2,923	24	11,825	18
<7 drinks/week	6,948	55	38,922	58
7+ drinks/week	1,277	10	8,750	13
Smoking				
Never smoked	6,080	49	34,342	52
Past smoker	5,313	43	28,348	43
Current smoker	1,056	8	3,852	5
Primary care provider				
Yes	2,439	93	63,999	95
Comorbidity index		0.56		0.38
Mean (SD)	12,623	(0.87)	67,368	(0.70)
Aspirin use				
Yes	2,676	21	14,817	22
Age at menarache, years				
<12	2,974	24	14,526	22
12–13	6,677	53	37,299	56
14+	2,904	23	15,320	22
Age at first birth, years				
Never pregnant/no term pregnancy	1,557	14	8,441	14
<20	1,827	16	7,146	12
20–29	6,895	62	40,369	66
30+	883	8	5,184	8

Table 2. Baseline Characteristics of Women with and without Depressive Symptoms: Women's Health Initiative Observational Cohort[a] (Cont'd)

	Depressive symptoms[b] n = *12,623*		*No depressive symptoms* n = *67,368*	
Variable	n	%	n	%
Number of children breastfed				
No term pregnancy	1,557	12	8,442	13
0	4,803	38	24,192	36
1–2	3,952	32	21,224	32
3+	2,147	18	12,778	19
Hormone therapy use				
Never used	5,038	40	26,349	39
Past user	1,818	14	9,139	14
Current user	5,767	46	31,880	47
First degree relatives with breast cancer				
None	9,549	84	52,010	84
1	1,573	14	8,643	14
2+	188	2	943	2
Breast reduction/removal				
Yes	634	5	2,877	4
History of benign breast disease				
No breast biopsy	9,659	78	51,417	77
One biopsy	1,913	15	10,374	16
2+ biopsies	849	7	4,547	7
Mammogram last year				
No mammogram ever	439	4	1,959	3
No mammogram last year	4,263	35	20,932	32
Mammogram last year	7,482	61	42,618	65
Ulcerative colitis or Crohn's disease				
Yes	215	2	724	1
First-degree relatives with colorectal cancer				
None	9,254	83	51,978	84
1	1,619	15	8,558	14
2+	239	2	1,021	2
Prior endoscopy				
None	5,787	46	31,184	47
<5 years ago	4,099	33	22,686	34
5+ years ago	2,553	21	12,493	19
Antidepressant use (at baseline)				
Yes	1,766	14	4,086	6
Antidepressant use (at 3-year follow-up)				
Yes	1,885	18	4,664	8
History of depression—DIS[c] (at baseline)				
Yes	5,152	42	4,427	7

[a]All analyses were statistically significant ($p < 0.0001$).
[b]Depressive symptoms: CES-D > 5 at baseline.
[c]DIS, Diagnostic Interview Schedule.

cer and 53% for colorectal cancer. Of women with current depressive symptoms at baseline, 61% reported screening for breast cancer compared with 65% of women without depressive symptoms. Differences in these rates persisted even when the rates were adjusted for factors associated with breast cancer screening (Table 3). Breast cancer screening rates during the average 7.6 years of follow-up among women with current depressive symptoms at baseline were 1.5 percentage points lower (−1.5 % difference, 95% CI −0.9, −2.0) compared with women without depressive symptoms, after adjustment for risk factors and differences between the two groups. Screening rates were even lower among women with a positive DIS at baseline (−3.6% difference, CI −2.9, −4.2) and among those with positive CES-D at both baseline and 3-year follow up (−2.2% difference, CI −1.3, −3.0) during the study.

Depressive symptoms at baseline or any other time during the follow-up were not associated with colorectal screening (Table 3). Neither breast cancer nor colorectal cancer stage at diagnosis (Table 4) was associated with current depressive symptoms, past depressive symptoms at baseline, or depressive symptoms at baseline and at the 3-year follow-up in unadjusted or adjusted analysis.

Table 3. Change in Cancer Screening with Depressive Symptoms: Women's Health Initiative Observational Cohort

	Breast cancer screening[a]		*Colorectal cancer screening*[b]	
Depressive symptoms	*Difference in percentage, %*[c]	*95% CI*	*Difference in percentage, %*[c]	*95% CI*
CES-D[d] > 5 at baseline	−1.5	−0.9, −2.0	0.2	1.1, −0.07
DIS > 2 at baseline	−3.6	−2.9, −4.2	0.0	1.0, 1.0
CES-D > 5 and DIS > 2 at baseline	−2.9	−2.0, −3.7	−0.1	1.3, −1.4
CES-D > 5 at baseline and 3-year follow-up	−2.2	−1.3, −3.0	0.6	2.0, −0.8

[a]Adjusted for sociodemographic characteristics, family history of breast cancer, history of previous breast biopsy, HT, alcohol intake, BMI, comorbidity index, insurance, and having a primary care provider.
[b]Adjusted for sociodemographic characteristics, family history of colorectal cancer, ulcerative colitis, or Crohn's disease; alcohol intake; comorbidity index; and having a primary care provider.
[c]Difference in percentage, %, comparing women with variable listed as reference group.
[d]CES-D, Center for Epidemiological Studies Depression scale; DIS: Diagnostic Interview Schedule.

Discussion

This study examined whether depressive symptoms are associated with breast and colorectal cancer screening rates and stage of diagnosis among postmenopausal women. To examine this hypothesis, we used WHI-OS data from 1991 to 1998, the largest (93,676) longitudinal study of healthy postmenopausal women. Among this large cohort, we found a high prevalence (15.8%) of women who screened positive for depressive symptoms at baseline, and almost half of them continued to have depressive symptoms at 3-year follow-up. Women with depressive symptoms at baseline had 1.5% lower breast cancer screening rates during the study period, controlling for other known predictors for screening. Breast cancer screening rates were even lower (−2.2% difference) among women reporting a past history of depression at baseline. This difference in breast cancer screening among women with depressive symptoms was not associated with presentation at a later stage of breast cancer. Depressive symptoms in the past, at baseline, or at 3-year follow-up had no association with adequate colorectal cancer screening rate or stage of colorectal cancer.

The incidence of cancer-related deaths in the year 2005 was the same as that in 1950.[22] Significant effort and resources have gone toward promoting early detection of cancer to reduce mortality and improve QOL. One possible barrier to early detection is co-morbid mental illness, such as depression. Numerous studies describe greater physical illness, functional impairment, and morbidity among patients with depression.[11,27–29] Patients with depression use more healthcare resources, including clinician visits and hospital admissions.[30,31] However, despite higher utilization rates, routine care such as screening may be overlooked. Patients and providers may be overly focused on managing the depression, or they may assume it is a natural consequence of events associated with aging, such as the loss of loved ones or medical illness, and appropriate proper treatment may not be considered. Additional barriers to seeking care include the social stigma associated with depression and social stressors, such as lack of social or financial support.[4]

Our results confirm that depressed women have modestly lower breast cancer screening rates, but we found no association with colorectal cancer screening rates. Pirraglia et al.[32] found similar results in a small cohort of younger women, aged 42–52 years, for whom high depressive symptom burden was a modest barrier for breast cancer screening but not for cervical cancer screening. Druss et al.[8,9] found that veterans with any mental illness or a dual diagnosis of mental illness and substance abuse were less likely to receive optimal cancer screening, including colorectal and prostate screening. Our study results differ from their studies for colorectal cancer screening. One possible reason for this differ-

Table 4. Odds Ratio of Later Stage of Cancer at Diagnosis: Women's Health Initiative Observational Cohort

Depressive symptoms	*Breast cancer*[a] *OR, (95% CI)*[c]	*Colorectal cancer*[b] *OR, (95% CI)*[c]
CES-D[d] > 5 at baseline	0.93 (0.71, 1.21)	1.01 (0.61, 1.69)
DIS > 2 at baseline	1.03 (0.77, 1.37)	1.1 (0.63, 1.92)
CES-D > 5 and DIS > 2 at baseline	0.97 (0.66, 1.44)	1.22 (0.59, 2.52)
CES-D > 5 at baseline and 3-year follow-up	1.17 (0.78, 1.74)	0.84 (0.34, 2.04)

[a]Adjusted for sociodemographic characteristics, insurance type, breast biopsy, number of relatives with breast cancer, moderate or strenuous physical activity, BMI, age at first birth, number of children breastfed, parity, and aspirin use.
[b]Adjusted for sociodemographic characteristics, insurance type, BMI, moderate or strenuous physical activity, and smoking status.
[c]OR, odds ratio; CI, confidence interval; CES-D, Center for Epidemiological Studies Depression scale; DIS: Diagnostic Interview Schedule.

ence may be that we examined the effect of depression independently from other mental illness, such as psychotic or anxiety disorders, which were combined in previous studies. As these mental illnesses differ significantly from depression in their severity impact on the patient, they may also differ in their relationship to obtaining cancer screening. WHI participants tended to be relatively healthy and self-motivated, as women with the most severe forms of mental illness were excluded from the study.

Although we found a moderate association of depressive symptoms and breast cancer screening, we found no association with colorectal cancer screening. Several possible reasons may account for this difference. Colorectal cancer screening can be adequate with less frequent screening modalities, such as endoscopy, which is repeated every 3–5 years. Because the severity of depressive symptoms can vary widely over time, it is possible that individuals may obtain screening during their symptom-free periods and thus achieve adequate colorectal cancer screening. On the other hand, breast cancer screening requires yearly mammography. A depressed woman may have difficulty adhering to annual appointments because of lack of interest in cancer screening, or she may have difficulty remembering appointments. Colorectal cancer screening is not as widely accepted as breast cancer screening. In our study, only 53% of all women received adequate colorectal cancer screening vs. 71% for breast cancer. Given the relatively lower rate of colorectal screening, the influence of depression may not be evident.

We did not find any association of depressive symptoms with breast or colorectal cancer in early stage or late stage at initial presentation. The results were consistent when all cancer stages were analyzed independently. One explanation for the lack of a finding may be that our cohort included fewer women with chronic, severe, or untreated depression. Second, this cohort may have had different baseline screening practices related to other factors, including their personal risk of cancer. The incidence rates of both breast and colorectal cancer were 3-fold and 1.5 fold higher, respectively, in the WHI-OS cohort than nationally reported by SEER data.[24] In the WHI-OS cohort, approximately 80% of the incident breast cancer cases diagnosed were in early stages compared with 60% nationally. Among colorectal cancer incident cases, 44% were diagnosed at the early stages, which is similar to national rates.

There are limitations to our study. The depression screening instrument (Burnam screen) may also reflect anxiety or psychological distress, and when it is used clinically, any patient with positive screening would generally be referred for further psychiatric evaluation. Although the positive screening result has high sensitivity and specificity for diagnosing clinical depression, it is only an indicator of depressive symptoms. The association between depressive symptoms and breast cancer screening might be stronger among subjects with severe or untreated depressive illness than is reported in this cohort of women. We were unable to verify the self-reported mammogram and colorectal cancer screening among these women. The literature supports high accuracy in self-reported screening data, 75% for breast[33] and 85% for colorectal screening,[34] if asked within 2 years. In the WHI study, women were queried about their health habits annually. Because of the limited data regarding the purpose of medication use, we could not control for antidepressant use and adherence to medications. The WHI was not designed to study the association of depressive symptoms with cancer stage.

One of the strengths of this study is the large number of participants, allowing us to adjust for most covariables without overfitting the multivariate models. Depression was measured at two different times during the study period using an instrument with reasonably high sensitivity and specificity for clinical depression. Sensitivity analysis by using a variety of definitions for depression, ranging from current depression to depression at multiple time points, was an added strength to the methodology. Additionally, stronger association of depressive symptoms with lower breast cancer screening rates among women with long-standing depressive symptoms also strengthened our results; we had enough power to study the effect of depressive symptoms on stage of cancer at presentation, which previous studies had been unable to do. Although our study is the first to study and show no effect of depressive symptoms on stage of breast and colorectal cancer, this association may be different among patients with severe refractory depression or other chronic mental illnesses. It is hoped that our study will help researchers to design additional studies looking at cancer screening and chronic mental illnesses.

In conclusion, we found that among a healthy and self-motivated cohort of women, self-reported depressive symptoms were associated with moderately lower rates of screening for breast cancer. No association was noted between depressive symptoms and adequacy of colorectal cancer screening. We were unable to find any association of depressive symptoms with stage of cancer at presentation. Cancer control programs might consider assessing psychiatric comorbidities, such as depression, when planning strategies to improve breast cancer screening rates.

Acknowledgments

Cordial thanks to R. Gold and colleagues for sharing the modified Charlson Index scores.

Women's Health Initiative Clinical Coordinating Center

Fred Hutchinson Cancer Research Center, Seattle, WA: Ross Prentice, Garnet Anderson, Andrea LaCroix, Charles L. Kooperberg, Ruth E. Patterson, Anne McTiernan; Wake Forest University School of Medicine, Winston-Salem, NC: Sally Shumaker; Medical Research Labs, Highland Heights, KY: Evan Stein; University of California at San Francisco, San Francisco, CA: Steven Cummings.

Women's Health Initiative WHI Clinical Centers

Albert Einstein College of Medicine, Bronx, NY: Sylvia Wassertheil-Smoller; Baylor College of Medicine, Houston, TX: Jennifer Hays; Brigham and Women's Hospital, Harvard Medical School, Boston, MA: JoAnn Manson; Brown University, Providence, RI: Annlouise R. Assaf; Emory University, Atlanta, GA: Lawrence Phillips; Fred Hutchinson Cancer Research Center, Seattle, WA: Shirley Beresford; George Washington University Medical Center, Washington, DC: Judith Hsia; Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance, CA: Rowan Chle-

bowski; Kaiser Permanente Center for Health Research, Portland, OR: Evelyn Whitlock; Kaiser Permanente Division of Research, Oakland, CA: Bette Caan; Medical College of Wisconsin, Milwaukee, WI: Jane Morley Kotchen; MedStar Research Institute/Howard University, Washington, DC: Barbara V. Howard; Northwestern University, Chicago/Evanston, IL: Linda Van Horn; Rush Medical Center, Chicago, IL: Henry Black; Stanford Prevention Research Center, Stanford, CA: Marcia L. Stefanick; State University of New York at Stony Brook, NY: Dorothy Lane; The Ohio State University, Columbus, OH: Rebecca Jackson; University of Alabama at Birmingham, Birmingham, AL: Cora E. Lewis; University of Arizona, Tucson/Phoenix, AZ: Tamsen Bassford; University at Buffalo, Buffalo, NY: Jean Wactawski-Wende; University of California at Davis, Sacramento, CA: John Robbins; University of California at Irvine, CA: F. Allan Hubbell; University of California at Los Angeles, CA: Howard Judd; University of California at San Diego, LaJolla/Chula Vista, CA: Robert D. Langer; University of Cincinnati, Cincinnati, OH: Margery Gass; University of Florida, Gainesville/Jacksonville, FL: Marian Limacher; University of Hawaii, Honolulu, HI: David Curb; University of Iowa, Iowa City/Davenport, IA: Robert Wallace; University of Massachusetts/Fallon Clinic, Worcester, MA: Judith Ockene; University of Medicine and Dentistry of New Jersey, Newark, NJ: Norman Lasser; University of Miami, FL: Mary Jo O'Sullivan; University of Minnesota, Minneapolis, MN: Karen Margolis; University of Nevada, Reno, NV: Robert Brunner; University of North Carolina, Chapel Hill, NC: Gerardo Heiss; University of Pittsburgh, Pittsburgh, PA: Lewis Kuller; University of Tennessee, Memphis, TN: Karen C. Johnson; University of Texas Health Science Center, San Antonio, TX: Robert Brzyski; University of Wisconsin, Madison, WI: Gloria E. Sarto; Wake Forest University School of Medicine, Winston-Salem, NC: Denise Bonds; Wayne State University School of Medicine/Hutzel Hospital, Detroit, MI: Susan Hendrix.

Disclosure Statement

No competing financial interests exist.

References

1. Jemal A, Murray T, Ward E, et al. Cancer statistics, 2005. CA Cancer J Clin 2005;55:10–30.
2. Seeff LC, Nadel MR, Klabunde CN, et al. Patterns and predictors of colorectal cancer test use in the adult U.S. population. Cancer 2004;100:2093–2103.
3. Husaini BA, Sherkat DE, Bragg R, et al. Predictors of breast cancer screening in a panel study of African American women. Women Health 2001;34:35–51.
4. NHMA. National Mental Health Association. Availabe at www.NMHA.org 2005.
5. Callahan CM, Wolinsky FD, Stump TE, Nienaber NA, Hui SL, Tierney WM. Mortality, symptoms, and functional impairment in late-life depression. J Gen Intern Med 1998; 13:746–752.
6. Callahan EJ, Bertakis KD, Azari R, Robbins JA, Helms LJ, Leigh JP. Association of higher costs with symptoms and diagnosis of depression. J Fam Pract 2002;51:540–544.
7. Hirschfeld RM, Keller MB, Panico S, et al. The National Depressive and Manic-Depressive Association consensus statement on the undertreatment of depression. JAMA 1997;277: 333–340.
8. Druss BG, Rosenheck RA, Desai MM, Perlin JB. Quality of preventive medical care for patients with mental disorders. Med Care 2002;40:129–136.
9. Desai MM, Rosenheck RA, Druss BG, Perlin JB. Receipt of nutrition and exercise counseling among medical outpatients with psychiatric and substance use disorders. J Gen Intern Med 2002;17:556–560.
10. Desai MM, Rosenheck RA, Druss BG, Perlin JB. Mental disorders and quality of diabetes care in the Veterans Health Administration. Am J Psychiatry 2002;159:1584–1590.
11. Ciechanowski PS, Katon WJ, Russo JE. Depression and diabetes: Impact of depressive symptoms on adherence, function, and costs. Arch Intern Med 2000;160:3278–3285.
12. Folsom DP, McCahill M, Bartels SJ, Lindamer LA, Ganiats TG, Jeste DV. Medical comorbidity and receipt of medical care by older homeless people with schizophrenia or depression. Psychiatr Serv 2002;53:1456–1460.
13. Levinson Miller C, Druss BG, Dombrowski EA, Rosenheck RA. Barriers to primary medical care among patients at a community mental health center. Psychiatr Serv 2003;54: 1158–1160.
14. Druss BG, Bradford WD, Rosenheck RA, Radford MJ, Krumholz HM. Quality of medical care and excess mortality in older patients with mental disorders. Arch Gen Psychiatry 2001;58:565–572.
15. 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.
16. Felker B, Yazel JJ, Short D. Mortality and medical comorbidity among psychiatric patients: A review. Psychiatr Serv 1996;47:1356–1363.
17. Design of the Women's Health Initiative clinical trial and observational study. The Women's Health Initiative Study Group. Control Clin Trials 1998;19:61–109.
18. Hjerl K, Andersen EW, Keiding N, Mouridsen HT, Mortensen PB, Jorgensen T. Depression as a prognostic factor for breast cancer mortality. Psychosomatics 2003;44: 24–30.
19. Burnam MA, Wells KB, Leake B, Landsverk J. Development of a brief screening instrument for detecting depressive disorders. Med Care 1988;26:775–789.
20. Robins L, Helzer J, Croughan J, Ratcliff K. National Institute of Mental Health Diagnostic Interview Schedule: Its history, characteristics, and validity. Arch Gen Psychiatry 1981;38: 381–389.
21. Tuunainen A, Langer RD, Klauber MR, Kripke DF. Short version of the CES-D (Burnam screen) for depression in reference to the structured psychiatric interview. Psychiatry Res 2001;103:261–270.
22. Americam Cancer Society. Available at *www.ACS.org. 2005 vol.*
23. Paley PJ. Screening for the major malignancies affecting women: Current guidelines. Am J Obstet Gynecol 2001; 184:1021–1030.
24. SEER. Surveillance Epidemiology and End Results. Last updated March 2006; accessed June 2006.
25. de Groot V, Beckerman H, Lankhorst GJ, Bouter LM. How to measure comorbidity. a critical review of available methods. J Clin Epidemiol 2003;56:221–229.
26. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudi-

nal studies: Development and validation. J Chronic Dis 1987;40:373–383.
27. Johnson J, Weissman MM, Klerman GL. Service utilization and social morbidity associated with depressive symptoms in the community. JAMA 1992;267:1478–1483.
28. Spitzer RL, Kroenke K, Linzer M, et al. Health-related quality of life in primary care patients with mental disorders. Results from the PRIME-MD 1000 Study. JAMA 1995;274: 1511–1517.
29. Bruce ML, Seeman TE, Merrill SS, Blazer DG. The impact of depressive symptomatology on physical disability: MacArthur Studies of Successful Aging. Am J Public Health 1994;84:1796–1799.
30. Borson S, Barnes RA, Kukull WA, et al. Symptomatic depression in elderly medical outpatients. I. Prevalence, demography, and health service utilization. J Am Geriatr Soc 1986;34:341–347.
31. Burns MJ, Cain VA, Husaini BA. Depression, service utilization, and treatment costs among Medicare elderly: Gender differences. Home Health Care Serv Q 2001;19:35–44.
32. Pirraglia PA, Rosen AB, Hermann RC, Olchanski NV, Neumann P. Cost-utility analysis studies of depression management: A systematic review. Am J Psychiatry 2004;161:2 155–2162.
33. Etzi S, Lane DS, Grimson R. The use of mammography vans by low-income women: the accuracy of self-reports. Am J Public Health 1994;84:107–109.
34. Madlensky L, McLaughlin J, Goel V. A comparison of self-reported colorectal cancer screening with medical records. Cancer Epidemiol Biomarkers Prev 2003;12:656–659.

Address reprint requests to:
Arpita Aggarwal, M.D., M.Sc.
Department of Internal Medicine
Virginia Commonwealth University
1200 East Broad Street
Richmond, VA 23298

E-mail: aaggarwal@mcvh-vcu.edu

Household Smoking Bans and Adolescent Antismoking Attitudes and Smoking Initiation: Findings From a Longitudinal Study of a Massachusetts Youth Cohort

Alison B. Albers, PhD, Lois Biener, PhD, Michael Siegel, MD, MPH, Debbie M. Cheng, ScD, and Nancy Rigotti, MD

Objectives. We sought to determine whether adolescents living in households in which smoking was banned were more likely to develop antismoking attitudes and less likely to progress to smoking compared with those living in households in which smoking was not banned.

Methods. We completed a longitudinal 4-year, 3-wave study of a representative sample of 3834 Massachusetts youths aged 12 to 17 years at baseline; 2791 (72.8%) were reinterviewed after 2 years, and 2217 (57.8%) were reinterviewed after 4 years. We used a 3-level hierarchical linear model to analyze the effect of a household ban on antismoking attitudes and smoking behaviors.

Results. The absence of a household smoking ban increased the odds that youths perceived a high prevalence of adult smoking, among both youths living with a smoker (odds ratio [OR]=1.56; 95% confidence interval [CI]=1.15, 2.13) and those living with nonsmokers (OR=1.75; 95% CI=1.29, 2.37). Among youths who lived with nonsmokers, those with no home ban were more likely to transition from nonsmoking to early experimentation (OR=1.89; 95% CI=1.30, 2.74) than were those with a ban.

Conclusions. Home smoking bans may promote antismoking attitudes among youths and reduce progression to smoking experimentation among youths who live with nonsmokers. (*Am J Public Health.* 2008;98:1886–1893. doi:10.2105/AJPH.2007.129320)

The proliferation of US smoke-free workplace policies and laws over the past decade has been accompanied by increased attention to private household smoking restrictions. The number of US households with comprehensive rules that make homes smoke free in all areas at all times has increased substantially.[1] The proportion of US households with smoke-free home rules increased from 43% in 1992 to 1993 to 72% in 2003.[2] Even smokers appear to be increasingly adopting such rules, particularly in homes in which they live with a nonsmoking adult.

Although smoke-free home bans are typically implemented to reduce or eliminate environmental tobacco smoke exposure in the household, these bans may have the additional benefit of reducing the initiation of smoking among youths by changing norms about the prevalence and social acceptability of smoking. Very little is known about the specific effect of a household smoking ban on youth smoking behavior or on smoking-related attitudes and norms that may mediate an effect on smoking behavior. In particular, few studies have addressed the independent effect of bans on youths who live with smokers—those who are at the greatest risk for becoming smokers themselves.

Recent studies showed that strong smoking regulations in local restaurants and bars were associated with more negative attitudes among youths toward the social acceptability of smoking in restaurants and bars.[3–6] Establishing household smoking bans conveys to youths living within these smoke-free home environments the message that smoking is unacceptable. Some supportive evidence, derived from cross-sectional data, indicates that a household smoking ban is associated with antismoking attitudes and norms. A recent cross-sectional study found that a household ban was associated with a lower perceived prevalence of adult smoking and more-negative attitudes about the social acceptability of smoking, 2 factors that affect the likelihood of smoking initiation.[7]

Several cross-sectional studies have reported that a smoking ban in the household was associated with a lower likelihood of being in an earlier stage of smoking and a lower current smoking prevalence among adolescents.[8–11] Conversely, other studies found no statistically significant association between a household smoking ban and reduced adolescent smoking.[12–14] Several factors may account for these conflicting results, including varying sample sizes, age groups, and smoking measures used in these cross-sectional studies.

A critical question is whether antismoking socialization occurs when parents themselves smoke. One study found that a household smoking ban was related to lower levels of smoking onset for children with nonsmoking parents but not for children with 1 or more parent who smoked.[15] Another study reported that a household smoking ban was not associated with trying smoking among high school students who had 1 or more parents who were current or former smokers.[16] Only 1 study reported an association between a household smoking ban and a reduced likelihood of smoking among 12th graders whose parents were smokers but not among those whose parents were nonsmokers.[17] In summary, more evidence supports an association between home smoking bans and lower levels of smoking behaviors among youths who live with nonsmokers.

Current research on household smoking bans has significant limitations. First, these studies rely on cross-sectional data that limit the ability to indicate causality in the relation between home smoking bans and trajectories of attitudes and smoking. Second, most studies have focused on individual-level predictors of attitudes and smoking behaviors, despite evidence that part of the explanation lies within the community context.[18] Third, few studies

have investigated the unique effects of a household smoking ban among adolescents living in home environments with parental smokers compared with those living with nonsmokers.

In this study, our goal was to improve existing research by (1) using longitudinal data that followed up a cohort of youths and young adults who lived in parental homes over a 4-year period, with a total of 3 repeated observations for each participant; (2) using a multilevel model that simultaneously examined the effects of individual-level and town-level factors; and (3) investigating separately the effects of a household ban on youths who live with at least 1 smoker and youths who live with nonsmokers.

METHODS

Design Overview

In this longitudinal 4-year, 3-wave study of a representative sample of 3834 Massachusetts youths aged 12 to 17 years at baseline (2001–2002), 2791 (72.8%) were reinterviewed after 2 years (in 2003–2004), and 2217 (57.8%) were reinterviewed after 4 years (in 2005–2006). We used a 3-level hierarchical linear model to analyze individual-level and town-level predictors of antismoking attitudes and smoking behaviors. The main predictor was a complete household smoking ban (yes vs no) assessed 2 years before measurement of the outcome. All analyses were stratified by youth report of family smoking: youths were categorized as living with a smoker if they responded that they have a parent or guardian who smokes cigarettes (1 or more smokers in household). In total, we examined 5 outcomes; the 3 attitudinal outcomes included (1) higher perceived prevalence of adult smoking in town, (2) perceived social acceptability of adult smoking in town, and (3) perceived social acceptability of youth smoking in town. The 2 behavioral outcomes included (1) the progression from nonsmoking to experimentation and (2) the transition from nonestablished to established smoking.

Sample

Between January 2001 and June 2002, the Center for Survey Research, University of Massachusetts, Boston, obtained a probability sample of 3834 Massachusetts youths, aged 12 to 17 years, by random-digit dialing. Between January 2003 and July 2004, we attempted to reinterview all youths in the baseline sample. Interviews were completed with 2791 youths, for a follow-up rate of 72.8%. Between January 2005 and July 2006, we attempted to reinterview all youths who responded to the wave-2 survey as well as those wave-1 youths who did not respond to wave 2 but for whom we had contact information. Interviews were completed with 2217 individuals (57.8% of the baseline sample). The analyses were restricted to youths who lived in parental homes, because they are primarily influenced by rules that have been established by other people.[10] Of those who completed wave 2, 88.9% (2481) lived with a parent at the time of interview, and of those who completed wave 3, 73.4% (1628) lived with a parent.

Measures and Outcome Variables

Complete household smoking ban. At waves 1 and 2, all youths were asked the question, "Some households have rules about when and where people may smoke. When you have visitors who smoke, are they allowed to smoke inside your home?" Youths who lived in a home in which at least 1 adult smoked were asked, "Do smokers in your household smoke inside your home?" Youths were categorized as having a complete household smoking ban if they reported that visitors were not allowed to smoke inside the home and, for those who lived in a home in which at least 1 adult smoked, if they reported a ban on smoking inside the home.

Perceived prevalence of adult smoking in town. Youths' perception of adult smoking prevalence in their town was based on their response to the following survey item: "About how many of the [TOWN] adults that you know smoke cigarettes?" Respondents who reported "very few" or "less than half" were classified as having a low level of perceived smoking prevalence for the adults in their town, whereas respondents who answered "about half," "more than half," or "almost all" were categorized as having a high level of perceived smoking prevalence for the adults in their town.

Social acceptability of adult and youth smoking in general in town. Two dichotomous measures of the perception of adult disapproval of smoking were assessed. Youth perception of adult disapproval of other adults smoking was based on the response to the following survey item: "How do most [TOWN] adults that you know feel about other adults smoking?" Youths were classified as perceiving adult smoking in general as "socially unacceptable" in their town if they responded that adults "disapprove a lot" or "disapprove a little," or "socially acceptable" in their town if they responded that adults "don't mind." Youth perception of adult disapproval of youths smoking was based on the response to the following item: "How do most [TOWN] adults that you know feel about teenagers smoking?" Youths were classified as perceiving youth smoking in general as "socially unacceptable" in their town if they responded that adults "disapprove a lot" or "disapprove a little" as opposed to "don't mind."

Stages of smoking initiation. Following Pierce et al.,[19] we defined progression to established smoking as having smoked 100 or more cigarettes. This measure has been formally validated[20–22] and used in previous studies.[20–26] We chose to use progression to established smoking as the sentinel measure of smoking initiation because it avoids the problem of the irregularity of smoking during adolescence.[22] It also avoids the problem of unreliable adolescent recall of smoking behavior during the previous 30 days by establishing a defined threshold of total lifetime cigarettes smoked to measure regular smoking behavior.

The experimentation stage of smoking was then defined as the period from trying a cigarette to becoming an established smoker. Thus, the stages of smoking initiation included (1) nonsmoking, (2) experimentation—having tried a cigarette but not smoked 100 cigarettes, and (3) established smoking—having smoked 100 or more cigarettes.

Individual-level covariates. We examined the effect of the following individual-level variables: (1) age group (12–14, 15–17, and 18–21 years), (2) gender, (3) race (non-Hispanic White vs other), (4) presence of at least 1 close friend who smokes, (5) education level of household informant (college graduate or not), (6) household income ($\leq$\$50 000 vs $>$\$50 000), (7) completed only wave-1 and wave-3 interviews, and (8) self-reported smoking status (nonsusceptible never smoker, susceptible

never smoker, puffer, experimenter, or current smoker).

Never smokers were defined as youths who had never puffed on a cigarette, puffers as those who had puffed but not smoked a whole cigarette, experimenters as those who had smoked at least 1 whole cigarette but none within the past 30 days, and current smokers as those who had smoked at least 1 cigarette, including 1 or more within the past 30 days. Never smokers were further classified as either susceptible or not susceptible to smoking on the basis of whether they indicated a firm commitment not to smoke in the future.[19,22,24,27,28]

In each attitudinal analysis, we controlled for attitudes at baseline of each transition period.

Town-level covariates. We examined the effect of the following town-level variables (included as continuous variables except when noted): (1) the percentage of each town's voters who voted "yes" on Question 1, a 1992 ballot initiative that increased the cigarette tax and created a statewide tobacco control program, (2) the percentage of White residents in each town, and (3) the percentage of youth (younger than 18 years) residents in each town. The percentage "yes" vote on Question 1 served as a measure of the baseline level of antismoking sentiment in each town before the proliferation of local restaurant smoking regulations, which has been found to correlate with the level of education in the town.[29] All town-level variables were obtained from the 2000 US Census, except for the data regarding the Question 1 vote, which was obtained from the Division of Elections within the Massachusetts Office of the Secretary of State.[29,30]

Data Analysis

This data set had clustering at 2 levels. First, observations were clustered within individual respondents. Each respondent could contribute up to 2 observations in the data set. Second, respondents were clustered within towns. Because observations among individuals and among respondents from the same town may be more similar than observations across respondents or across respondents from different towns, we used a multilevel (hierarchical) logistic regression model to examine the relation between absence of a household smoking ban at baseline and the study

TABLE 1—Baseline Characteristics of Cohort and Presence of a Complete Smoking Ban in Household, by Individual and Contextual Variables Among Youths Living With Smokers or Those Living With Nonsmokers: Massachusetts, 2001–2006

	Lived With Smoker		Lived Only With Nonsmokers	
	Household Smoking Ban	No Household Smoking Ban	Household Smoking Ban	No Household Smoking Ban
Total, no. (%)	724 (100)	672 (100)	2276 (100)	277 (100)
Attitudinal outcome variables				
Perceived prevalence of adult smoking in town, no. (%)				
Low (fewer than half)	391 (54.0)	239 (36.0)	1703 (74.9)	183 (65.5)
High (half or more)	329 (46.0)	425 (64.0)	566 (25.1)	92 (34.5)
Social acceptability of smoking by adults in town, no. (%)				
Unacceptable	410 (54.5)	271 (42.1)	1646 (73.2)	148 (53.7)
Acceptable	311 (45.6)	395 (57.9)	615 (26.8)	129 (46.3)
Social acceptability of smoking by youths in town, no. (%)				
Unacceptable	672 (92.9)	599 (90.0)	2177 (96.1)	257 (93.3)
Acceptable	49 (7.1)	66 (10.0)	91 (3.9)	20 (6.7)
Behavioral outcome variables				
Progression to established smoking, no. (%)				
No	578 (87.5)	490 (85.8)	2053 (93.3)	233 (91.3)
Yes	81 (12.5)	84 (14.2)	147 (6.7)	20 (6.7)
Progression from nonsmoking to experimental smoking, no. (%)				
No	434 (77.9)	369 (76.8)	1722 (84.9)	175 (78.9)
Yes	122 (22.1)	114 (23.2)	299 (15.1)	47 (21.1)
Individual-level time-varying covariates (level 1)				
Age group, y, no. (%)				
12–14	294 (38.9)	274 (40.5)	936 (41.3)	116 (40.2)
15–17	373 (52.8)	353 (52.7)	1196 (52.0)	148 (55.0)
18–21	57 (8.3)	45 (6.8)	144 (6.7)	13 (4.8)
Baseline smoking status, no. (%)				
Nonsusceptible never smoker	203 (28.7)	203 (31.1)	956 (42.0)	104 (38.0)
Susceptible never smoker	251 (33.8)	183 (27.5)	882 (39.1)	98 (35.5)
Puffed	102 (13.9)	97 (13.4)	184 (8.3)	20 (7.2)
Smoked whole cigarette	79 (10.8)	72 (11.1)	147 (6.1)	28 (9.6)
Smoked in past 30 d	89 (12.8)	117 (16.9)	107 (4.5)	27 (9.7)
Presence of close friend who smokes, no. (%)				
No	428 (59.0)	354 (53.4)	1680 (74.2)	196 (70.5)
Yes	295 (41.0)	318 (46.6)	595 (25.8)	80 (29.5)
Participated in 4-y follow-up (wave 1 to wave 3), no. (%)				
No	700 (96.7)	640 (95.6)	2221 (97.7)	266 (96.0)
Yes	24 (3.3)	32 (4.4)	55 (2.3)	11 (4.0)
Individual-level covariates (level 2)				
Gender				
Boy	370 (50.9)	327 (50.3)	1173 (51.5)	164 (59.6)
Girl	354 (49.1)	345 (49.7)	1103 (48.5)	113 (40.4)
Race/ethnicity				
Non-Hispanic White	570 (79.8)	555 (82.7)	1862 (82.3)	244 (88.3)
Other	153 (20.2)	110 (17.3)	397 (17.8)	31 (11.8)

Continued

TABLE 1—*Continued*

Household income, $, no. (%)				
≤50 000	222 (37.9)	246 (44.7)	357 (18.7)	55 (23.6)
>50 000	343 (62.1)	297 (55.3)	1532 (81.3)	189 (76.4)
Informant education, no. (%)				
Not college graduate	491 (68.1)	476 (73.0)	1008 (45.0)	160 (58.3)
College graduate	219 (31.9)	175 (27.0)	1234 (55.0)	113 (41.8)
Town-level covariates (level 3)				
"Yes" vote on Question 1,[a] mean (%)	724 (49.6)	672 (47.9)	2276 (51.9)	277 (49.9)
White residents, mean (%)	724 (88.0)	672 (86.6)	2276 (88.4)	277 (89.2)
Residents who are youths, mean (%)	724 (24.7)	672 (24.4)	2276 (24.7)	277 (24.7)

Note. Table entries are weighted percentages.
[a]Question 1 was a 1992 ballot initiative that increased the cigarette tax and created a statewide tobacco control program.

outcomes. This procedure accounts for correlation of data within individuals and within town "clusters," reducing the probability of a type-I error that could be introduced if this correlation were ignored.[31,32]

All town-level variables were time-independent and assessed at the start of the study (modeled at level 3). Time-independent individual-level covariates (entered at level 2) were gender, race, informant education, and household income. The following individual-level covariates could change from survey to survey and were modeled at level 1: the presence of a household smoking ban, age group, presence of a close friend who smokes, and absence of a household smoking ban at baseline. All analyses were stratified by the time-varying covariate of living with at least 1 smoker in the household.

For the baseline sample, survey weights were computed that adjusted for the number of telephones per household and hence for the probability of selection. We made adjustments to the baseline weights with a raking procedure to correct for biased attrition. All analyses were conducted with 2-sided tests and a significance level of .05. Analyses were conducted with HLM 6.0 (Scientific Software International Inc, Lincolnwood, IL).

RESULTS

Baseline Characteristics of Sample

Our study sample consisted of 2593 unique individuals who reported a smoking ban in their household and lived with a parent, contributing a total of 3949 observations. Of the 3949 observations, 1396 (35.4%) were from youths who lived with at least 1 smoker in the household, and 2553 (64.6%) were from those who lived with nonsmokers. Of the 1396 observations from youths who lived with a smoker, 51.9% reported a complete household smoking ban; of the total 2553 observations from youths who lived with nonsmokers, 89.2% reported a complete household smoking ban (Table 1).

Attitudinal Outcomes

For the multivariate analyses of the attitudinal outcomes, our study sample consisted of 942 unique participants (contributing a total of 1394 observations) who lived with a smoker and 1728 unique participants (contributing a total of 2634 observations) who did not live with a smoker.

Predictors of perceived smoking prevalence. Youths living in households that lacked a complete household smoking ban were more likely to perceive a high prevalence of adult smoking in their town compared with youths who lived in households with a complete smoking ban (Table 2). The relation existed both for youths who lived with a smoker (odds ratio [OR]=1.56; 95% confidence interval [CI]=1.15, 2.13) and for youths who lived with nonsmokers (OR=1.75; 95% CI=1.29, 2.37).

Predictors of perceived social acceptability of adult and youth smoking. Youths living in a household without a household smoking ban also were more likely to consider adult smoking to be socially acceptable than were youths who lived in homes with a smoking ban. The magnitude of effect of the household smoking ban was similar for youths who lived with a smoker (OR=1.55; 95% CI=1.21, 1.99) and for those who did not live with a smoker (OR=1.53; 95% CI=1.26, 2.22; Table 2).

A clinically important, although not statistically significant, effect of a complete household smoking ban was seen on the perceived social acceptability of youth smoking for youths who lived with a smoker (OR=1.66 for absence of household smoking ban vs presence of household smoking ban; 95% CI=0.93, 2.98). However, a complete household smoking ban had no effect on perceived youth smoking prevalence in town among youths who lived with nonsmokers (OR=1.04; 95% CI=0.58, 1.89; Table 2).

Behavioral Outcomes

For analysis of the behavioral outcomes, overall progression to established smoking for youths who lived with a smoker was based on 1241 observations (wave 1 to wave 2: 738; wave 2 to wave 3: 451; wave 1 to wave 3: 52), and analysis of overall progression to established smoking for youths who did not live with a smoker was based on 2541 observations (wave 1 to wave 2: 1604; wave 2 to wave 3: 872; wave 1 to wave 3: 65). Analyses of progression from nonsmoking to experimentation for youths who lived with a smoker were based on a total of 1042 observations (wave 1 to wave 2: 631; wave 2 to wave 3: 370; wave 1 to wave 3: 41), and analyses of progression to experimentation for youths who did not live with a smoker were based on 2318 observations (wave 1 to wave 2: 1469; wave 2 to wave 3: 784; wave 1 to wave 3: 65). Sample sizes for the analyses of the transition from nonsmoking to experimentation were slightly smaller than for overall progression to established smoking because transition from nonsmoking to experimentation included only those observations for nonsmoking youths at baseline of the transition period.

Predictors of overall progression to established smoking. The lack of a complete household smoking ban had no effect on progression to established smoking, either for youths who lived with a smoker (OR=1.38; 95% CI=0.92, 2.07) or for youths who lived with nonsmokers (OR=1.08; 95% CI=0.61, 1.93; Table 3).

TABLE 2—Adjusted Odds Ratios (ORs) for Perceived Smoking Prevalence and Acceptability of Smoking Among Youths Living With a Smoker and Those Living With Nonsmokers: Massachusetts, 2001–2006

	Higher Perceived Prevalence of Adult Smoking in Town[a]		Social Acceptability of Smoking by Adults in Town[b]		Social Acceptability of Smoking by Youths in Town[c]	
	Lived With Smoker, OR (95% CI)	Lived With Nonsmokers, OR (95% CI)	Lived With Smoker, OR (95% CI)	Lived With Nonsmokers, OR (95% CI)	Lived With Smoker, OR (95% CI)	Lived With Nonsmokers, OR (95% CI)
Main time-varying predictor variable (level 1)						
Presence of a complete smoking ban in household						
Yes (Ref)	1.00	1.00	1.00	1.00	1.00	1.00
No	1.56* (1.15, 2.13)	1.75* (1.29, 2.37)	1.55* (1.21, 1.99)	1.53* (1.26, 2.22)	1.66 (0.93, 2.98)	1.04 (0.58, 1.89)
Individual-level time-varying covariates (level 1)						
Age group, y						
12–14 (Ref)	1.00	1.00	1.00	1.00	1.00	1.00
15–17	0.76** (0.59, 0.99)	0.93 (0.74, 1.18)	0.93 (0.71, 1.23)	1.09 (0.88, 1.35)	1.59** (0.97, 2.60)	2.20* (1.40, 3.48)
18–21	0.65 (0.40, 1.05)	0.66 (0.41, 1.09)	0.97 (0.59, 1.60)	0.86 (0.76, 1.73)	0.81 (0.37, 1.92)	0.82 (0.45, 2.56)
Self-reported baseline smoking status						
Nonsusceptible never smoker (Ref)	1.00	1.00	1.00	1.00	1.00	1.00
Susceptible never smoker	0.92 (0.68, 1.26)	0.92 (0.74, 1.15)	0.86 (0.61, 1.21)	1.12 (0.89, 1.42)	0.84 (0.42, 1.68)	0.75 (0.42, 1.33)
Puffed	1.02 (0.60, 1.75)	1.04 (0.72, 1.51)	1.29 (0.88, 1.91)	1.34 (0.84, 2.15)	1.37 (0.63, 2.98)	0.65 (0.25, 1.67)
Smoked whole cigarette	0.91 (0.57, 1.47)	1.53 (0.98, 2.41)	1.15 (0.73, 1.80)	1.04 (0.67, 1.63)	2.21 (0.99, 4.94)	0.87 (0.35, 2.18)
Smoked in past 30 d	1.43 (0.84, 2.45)	1.47 (0.73, 2.96)	1.45 (0.85, 2.48)	1.54 (0.99, 2.38)	1.86 (0.88, 3.93)	1.64 (0.72, 3.74)
Presence of close friend who smokes						
No (Ref)	1.00	1.00	1.00	1.00	1.00	1.00
Yes	1.63* (1.15, 2.30)	1.06 (0.79, 1.42)	1.25 (0.92, 1.69)	1.44* (1.13, 1.83)	1.18 (0.70, 2.01)	1.38 (0.84, 2.29)
Participated in 4-y follow-up (wave 1 to wave 3)						
No (Ref)	1.00	1.00	1.00	1.00	1.00	1.00
Yes	1.58 (0.73, 3.41)	0.76 (0.40, 1.44)	0.81 (0.43, 1.51)	1.10 (0.55, 2.18)	2.59** (0.89, 5.22)	3.52* (1.53, 8.12)
Baseline attitude	4.09* (2.98, 5.64)	4.14* (3.26, 5.27)	2.78* (2.03, 3.80)	3.06* (2.41, 3.90)	1.64 (0.77, 3.50)	3.37* (1.48, 7.70)
Individual-level covariates (level 2)						
Gender						
Boy (Ref)	1.00	1.00	1.00	1.00	1.00	1.00
Girl	0.81 (0.59, 1.10)	0.97 (0.77, 1.23)	0.94 (0.74, 1.20)	0.79** (0.64, 0.97)	0.84 (0.56, 1.25)	0.59** (0.39, 0.91)
Race/ethnicity						
Non-Hispanic White (Ref)	1.00	1.00	1.00	1.00	1.00	1.00
Other	1.60** (1.07, 2.38)	1.50** (1.06, 2.11)	0.92 (0.63, 1.35)	1.61* (1.18, 2.20)	1.64 (0.97, 2.76)	1.60** (0.95, 2.72)
Household income, $						
≤ 50 000 (Ref)	1.00	1.00	1.00	1.00	1.00	1.00
> 50 000	0.61* (0.44, 0.84)	0.62* (0.45, 0.86)	0.88 (0.64, 1.22)	1.10 (0.85, 1.44)	1.16 (0.74, 1.81)	0.87 (0.51, 1.50)
Informant education						
Not college graduate (Ref)	1.00	1.00	1.00	1.00	1.00	1.00
College graduate	0.57* (0.41, 0.80)	0.90 (0.68, 1.20)	0.83 (0.60, 1.15)	0.90 (0.72, 1.12)	0.68 (0.43, 1.08)	0.76 (0.48, 1.18)
Town-level covariates (level 3)[d]						
Percentage "yes" vote on Question 1[e]	0.55* (0.45, 0.67)	0.55* (0.46, 0.64)	0.85** (0.70, 1.03)	0.72* (0.63, 0.81)	0.96 (0.67, 1.35)	0.72* (0.53, 0.96)
Percentage of residents who are White	0.96 (0.86, 1.08)	1.03 (0.93, 1.13)	0.87** (0.78, 0.97)	1.12** (1.01, 1.24)	0.99 (0.84, 1.19)	1.15 (0.92, 1.44)
Percentage of residents who are youths	1.05 (0.71, 1.54)	0.78 (0.56, 1.07)	1.11 (0.78, 1.58)	0.95 (0.72, 1.27)	0.61 (0.34, 1.10)	0.38 (0.21, 0.69)

Note. CI = confidence interval.

[a]Perceived prevalence of adult smoking in town was coded as 0 (lower perception of smoking prevalence; reference category) and 1 (higher perception of smoking prevalence). Analyses were on the basis of 942 individuals living in 234 towns, contributing a total of 1391 observations for youths living with smokers, and on 1728 individuals living in 280 towns, contributing a total of 2631 observations for youths living without smokers.

[b]Social acceptability of smoking by adults in town was coded as 0 (perceived adult disapproval of adult smoking; reference category) and 1 (no perceived disapproval of adult smoking). Analyses were on the basis of 942 individuals living in 234 towns, contributing a total of 1394 observations for youths living with smokers, and on 1720 individuals living in 280 towns, contributing a total of 2619 observations for youths living without smokers.

[c]Social acceptability of smoking by youths in town was coded as 0 (perceived adult disapproval of youth smoking; reference category) and 1 (no perceived disapproval of youth smoking). Analyses on the basis of 941 individuals living in 234 towns, contributing a total of 1393 observations for youths living with smokers and analyses on the basis of 1725 individuals living in 280 towns, contributing a total of 2628 observations for youths living without smokers.

TABLE 3—Adjusted Odds Ratios (ORs) for Overall Progression to Established Smoking and Transition From Nonsmoking to Experimentation Among Youths Living With a Smoker and Those Living With Nonsmokers: Massachusetts, 2001–2006

	Overall Progression to Established Smoking[a]		Transition From Nonsmoking to Experimentation[b]	
	Lived With Smoker, OR (95% CI)	Lived With Nonsmokers, OR (95% CI)	Lived With Smoker, OR (95% CI)	Lived With Nonsmokers, OR (95% CI)
Main predictor variable (level 1)				
Presence of a complete smoking ban in household				
Yes (Ref)	1.00	1.00	1.00	1.00
No	1.38 (0.92, 2.07)	1.08 (0.61, 1.93)	0.99 (0.73, 1.37)	1.89* (1.30, 2.74)
Individual-level time-varying covariates (level 1)				
Age group, y				
12–14	1.00	1.00	1.00	1.00
15–17	0.83 (0.52, 1.31)	1.72** (1.11, 2.65)	0.93 (0.86, 1.02)	2.20* (1.65, 2.93)
18–21	0.69 (0.32, 1.49)	0.86 (0.44, 1.67)	0.98 (0.35, 2.71)	1.22 (0.69, 2.17)
Baseline smoking status				
Nonsusceptible never smoker	1.00	1.00	1.00	1.00
Susceptible never smoker	1.43 (0.72, 2.85)	1.96** (1.10, 3.48)	0.92 (0.64, 1.30)	1.24 (0.93, 1.67)
Puffed	5.51* (2.71, 11.20)	4.95* (2.27, 10.82)	. . .	. . .
Smoked whole cigarette	12.95* (6.03, 27.77)	19.41* (9.92, 37.99)	. . .	. . .
Smoked in past 30 d	43.14* (17.35, 107.3)	49.08* (23.20, 103.8)	. . .	. . .
Presence of close friend who smokes				
No	1.00	1.00	1.00	1.00
Yes	1.90* (1.27, 2.84)	2.34* (1.08, 2.52)	1.93* (1.38, 2.70)	2.26* (1.64, 3.12)
4-y follow-up period (wave 1 to wave 3)				
No	1.00	1.00	1.00	1.00
Yes	3.78* (1.81, 7.85)	4.17* (1.28, 13.59)	1.89 (0.77, 4.65)	2.81* (1.39, 5.67)
Individual-level covariates (level 2)				
Gender				
Boy (Ref)	1.00	1.00	1.00	1.00
Girl	0.72 (0.48, 1.09)	0.43* (0.29, 0.63)	0.90 (0.65, 1.26)	0.80 (0.63, 1.03)
Race/ethnicity				
Non-Hispanic White (Ref)	1.00	1.00	1.00	1.00
Other	1.17 (0.62, 2.21)	1.24 (0.69, 2.25)	0.98 (0.61, 1.57)	0.80 (0.47, 1.36)
Household income, $				
≤50 000 (Ref)	1.00	1.00	1.00	1.00
>50 000	0.81 (0.49, 1.31)	1.46 (0.82, 2.60)	1.14 (0.76, 1.72)	1.72** (1.10, 2.68)
Informant education				
Not college graduate (Ref)	1.00	1.00	1.00	1.00
College graduate	0.79 (0.48, 1.32)	1.08 (0.69, 1.70)	1.18 (0.80, 1.72)	0.95 (0.69, 1.31)
Town-level covariates (level 3)[c]				
Percentage "yes" vote on Question 1[d]	0.89 (0.68, 1.17)	1.24 (0.96, 1.59)	1.01 (0.83, 1.24)	1.09 (0.92, 1.30)
Percentage of residents who are White	1.31* (1.08, 1.58)	1.12 (0.88, 1.44)	1.15 (0.96, 1.36)	1.01 (0.88, 1.18)
Percentage of residents who are youths	0.77 (0.40, 1.46)	0.86 (0.53, 1.41)	0.79 (0.51, 1.24)	0.88 (0.65, 1.21)

Note. CI = confidence interval. Ellipses indicate not applicable.
[a]Progression to established smoking is defined as having smoked 100 cigarettes in one's lifetime. Analyses on the basis of 858 individuals living in 229 towns, contributing a total of 1241 observations for youths living with smokers, and analyses on the basis of 1672 individuals living in 276 towns, contributing a total of 2538 observations for youths living without smokers.
[b]Analyses on the basis of 731 individuals living in 211 towns, contributing a total of 1042 observations for youths living with smokers, and analyses on the basis of 1538 individuals living in 268 towns, contributing a total of 2315 observations for youths living without smokers.
[c]Odds ratio associated with each 10-percentage-point increase in variable.
[d]Question 1 was a 1992 ballot initiative that increased the cigarette tax and created a statewide tobacco control program.
*$P < .01$; **$P < .05$.

Predictors of transition from nonsmoking to experimentation. Among youths who lived with nonsmokers, the absence of a complete household smoking ban increased the odds of transitioning from nonsmoking to experimentation (OR=1.89; 95% CI=1.30, 2.74; Table 3). However, the absence of a complete household smoking ban had no effect on the transition from nonsmoking to experimentation among those who lived with a smoker (OR=0.99; 95% CI=0.73, 1.37; Table 3).

DISCUSSION

To the best of our knowledge, this was the first longitudinal study to examine the effects of household smoking bans on adolescents' attitudes about the acceptability of smoking, perceptions of smoking prevalence, and likelihood of initiating smoking and to assess these relations separately among youths who lived with a smoker and those who lived with nonsmokers. We used a hierarchical, repeated measures model and found that, among youths who did and did not live with a smoker, having a home smoking ban significantly increased the odds that adolescents would have negative attitudes about the social acceptability of smoking. Having a home smoking ban reduced the odds that an adolescent would begin to experiment with cigarettes but only in homes that did not contain smokers. The presence of a household smoking ban did not reduce progression to established smoking, regardless of whether a smoker lived in the home.

These findings provide support for the hypothesis that household smoking bans provide parents with an antismoking measure that contributes to antitobacco socialization of their children.[10] Even in the presence of parental smoking, prohibiting smoking in the home and clearly communicating household smoking rules may lower youths' perception of smoking prevalence and attitudes about the social acceptability of smoking.

We found no effect of a complete household smoking ban on overall progression to established smoking among adolescents. This result was unexpected, but it is consistent with Fisher et al.'s recent study findings.[14] Because parental disapproval and negative parental attitudes toward smoking have been shown to decrease the likelihood of adolescent smoking, we expected a household smoking ban to reduce progression to established smoking, but it did not. We examined several alternative measures of smoking initiation (data not shown) and found no effect of household smoking bans on smoking initiation, regardless of whether a youth lived in a household with a smoker.

We did find that a household smoking ban reduced early experimentation with cigarettes but only among youths who lived with nonsmokers. Reducing experimentation with smoking may require that youths live in a home with a consistent message—nonsmoking parents who ban smoking entirely in the home. Household antismoking restrictions may be effective only when they match parental behavior.[16,33]

These results have several important public health policy implications. First, this study found an effect of a complete household smoking ban on perceived smoking prevalence and the perceived social acceptability of smoking above and beyond a host of individual- and town-level predictors. Thus, household smoking bans may be an effective intervention to promote antismoking attitudes, particularly among those who are at the greatest risk for exposure to smoking.

Second, these findings confirm the potential moderating effect of parental smoking. Contrary to our expectations, we found an effect of a complete household ban on the transition from nonsmoking to early experimentation only among those who lived with nonsmokers. The fact that a family member smokes may be a more important determinant of adolescent smoking than are the arrangements made to restrict smoking in the household.[12]

The primary potential threat to the validity of our findings is the relatively high rate of loss to follow-up in the study. Although not unusual for a telephone survey that followed up participants for 4 years, the follow-up rates of 73% at wave 2 and 58% at wave 3 do introduce the possibility of a differential loss to follow-up bias.[18] Analyses of the baseline differences between youth respondents to either wave 2 or wave 3 and those who failed to respond did indicate that responders were significantly more likely to report having a home smoking ban. To help correct for biased attrition, we used variables that were significantly related to having a smoking ban (including parental education and youth smoking status) in an iterative raking procedure to create adjustments to the baseline weights.

A second limitation of this research is that it relied on youth report of home smoking policies. Although youth reports may be less accurate than parental reports, they measure youth perception, which may be more important than actual household restrictions.[12]

The evidence presented in this article supports the conclusion that the presence of complete household smoking bans significantly increases the likelihood that youths will develop antismoking attitudes and decreases the likelihood of youth smoking initiation by impeding the progression from nonsmoking to early cigarette experimentation among youths living with nonsmokers. This study supports the notion that home smoking bans have the potential to promote antismoking norms and to prevent adolescent smoking. ■

About the Authors

Alison B. Albers and Michael Siegel are with the Social and Behavioral Sciences Department, Boston University School of Public Health, Boston, MA. Lois Biener is with the Center for Survey Research, University of Massachusetts, Boston. Debbie M. Cheng is with the Biostatistics Department, Boston University School of Public Health, Boston. Nancy Rigotti is with Massachusetts General Hospital, Boston, and Harvard Medical School, Boston.

Requests for reprints should be sent to Alison B. Albers, PhD, Social and Behavioral Sciences Department, Boston University School of Public Health, 801 Massachusetts Ave, Crosstown Center 4th Floor, Boston, MA 02118 (e-mail: aalbers@bu.edu).

This article was accepted February 15, 2008.

Contributors

A. B. Albers conducted the data analyses and wrote an initial draft of the article. L. Biener was the principal investigator of the parent study, directed survey administration and data collection, interpreted study findings, and edited drafts of the article. M. Siegel interpreted study findings and edited drafts of the article. D. M. Cheng was responsible for analytic design, statistical and methodological guidance, and data interpretation. N. Rigotti is the principal investigator of the grant that funded this study and originated the idea for the study, interpreted study findings, and edited drafts of the article. All of the authors contributed to conceptualization of the research question and the design of the study and reviewed and edited the final version of the article.

Acknowledgments

This work was supported by the Flight Attendant Medical Research Institute and the National Cancer Institute's State and Community Tobacco Control Interventions Research Grant Program (grant 2R01 CA86257).

Human Participant Protection

This study was approved by the institutional review boards of the University of Massachusetts Boston (survey administration and data collection site) and the Boston University Medical Center (data analysis site for the study described in this article).

References

1. US Department of Health and Human Services. *The Health Consequences of Involuntary Exposure to Tobacco Smoke: A Report of the Surgeon General.* Atlanta, Ga: Office on Smoking and Health, Coordinating Center for Health Promotion, National Center for Chronic Disease Prevention and Health Promotion; 2006.

2. Centers for Disease Control and Prevention. Exposure to secondhand smoke among students aged 13–15 years—worldwide, 2000–2007. *MMWR Morb Mortal Wkly Rep.* 2007;56:497–500.

3. Siegel M, Albers AB, Cheng DM, Biener L, Rigotti NA. Effect of local restaurant smoking regulations on environmental tobacco smoke exposure among youths. *Am J Public Health.* 2004;94:321–325.

4. Albers AB, Siegel M, Cheng DM, Rigotti NA, Biener L. Effects of restaurant and bar smoking regulations on exposure to environmental tobacco smoke among Massachusetts adults. *Am J Public Health.* 2004;94:1959–1964.

5. Siegel M, Albers AB, Cheng DM, Biener L, Rigotti NA. Effect of local restaurant smoking regulations on progression to established smoking among youths. *Tob Control.* 2005;14:300–306.

6. Albers AB, Siegel M, Cheng DM, Biener L, Rigotti NA. Relation between local restaurant smoking regulations and attitudes towards the prevalence and social acceptability of smoking: a study of youths and adults who eat out predominantly at restaurants in their town. *Tob Control.* 2004;13:347–355.

7. Conley Thomson C, Siegel M, Winickoff J, Biener L, Rigotti NA. Household smoking bans and adolescents' perceived prevalence of smoking and social acceptability of smoking. *Prev Med.* 2005;41:349–356.

8. Farkas AJ, Gilpin EA, White MM, Pierce JP. Association between household and workplace smoking restrictions and adolescent smoking. *JAMA.* 2000;284:717–722.

9. Wakefield MA, Chaloupka FJ, Kaufman NJ, et al. Effect of restrictions on smoking at home, at school, and in public places on teenage smoking: cross sectional study. *BMJ.* 2000;321:333–337.

10. Clark PI, Schooley MW, Pierce B, Schulman J, Hartman AM, Schmitt CL. Impact of home smoking rules on smoking patterns among adolescents and young adults. *Prev Chronic Dis.* 2006;3(2):A41.

11. Henriksen L, Jackson C. Anti-smoking socialization: relationship to parent and child smoking status. *Health Commun.* 1998;10:87–101.

12. Biener L, Cullen D, Di ZX, Hammond SK. Household smoking restrictions and adolescents' exposure to environmental tobacco smoke. *Prev Med.* 1997;26:358–363.

13. Harakeh Z, Scholte RH, de Vries H, Engels RC. Parental rules and communication: their association with adolescent smoking. *Addiction.* 2005;100:862–870.

14. Fisher LB, Winickoff JP, Camargo CA Jr, Colditz GA, Frazier AL. Household smoking restrictions and adolescent smoking. *Am J Health Promot.* 2007;22:15–21.

15. Jackson C, Henriksen L. Do as I say: parent smoking, antismoking socialization and smoking onset among children. *Addict Behav.* 1997;22:107–114.

16. Proescholdbell RJ, Chassin L, MacKinnon DP. Home smoking restrictions and adolescent smoking. *Nicotine Tob Res.* 2000;2:159–167.

17. Andersen MR, Leroux BG, Bricker JB, Rajan KB, Peterson AV Jr. Antismoking parenting practices are associated with reduced rates of adolescent smoking. *Arch Pediatr Adolesc Med.* 2004;158:348–352.

18. Siegel M, Albers AB, Cheng DM, Hamilton WL, Biener L. Local restaurant smoking regulations and the adolescent smoking initiation process: results of a multilevel contextual analysis among Massachusetts youths. *Arch Pediatr Adolesc Med.* In press.

19. Pierce JP, Choi WS, Gilpin EA, et al. Tobacco industry promotion of cigarettes and adolescent smoking. *JAMA.* 1998;279:511–515.

20. Pierce JP, Choi WS, Gilpin EA, Farkas AJ, Merritt RK. Validation of susceptibility as a predictor of which adolescents take up smoking in the United States. *Health Psychol.* 1996;15:355–361.

21. Choi WS, Pierce JP, Gilpin EA, Farkas AJ, Berry CC. Which adolescent experimenters progress to established smoking in the United States. *Am J Prev Med.* 1997;13:385–391.

22. Pierce JP, Farkas AJ, Evans N, et al. An improved surveillance measure for adolescent smoking? *Tob Control.* 1995;4(suppl 1):S47–S56.

23. Mowery PD, Farrelly MC, Haviland L, Gable JM, Wells HE. Progression to established smoking among US youths. *Am J Public Health.* 2004;94:331–337.

24. Siegel M, Biener L. The impact of an antismoking media campaign on progression to established smoking: results of a longitudinal youth study. *Am J Public Health.* 2000;90:380–386.

25. Siegel M, Biener L, Rigotti NA. The effect of local tobacco sales laws on adolescent smoking initiation. *Prev Med.* 1999;29:334–342.

26. Choi WS, Ahluwalia JS, Harris KJ, et al. Progression to established smoking: the influence of tobacco marketing. *Am J Prev Med.* 2002;22:228–233.

27. Pierce JP, Choi WS, Gilpin EA, et al. Validation of susceptibility as a predictor of which adolescents take up smoking in the United States. *Health Psychol.* 1996;15:355–361.

28. Choi WS, Pierce JP, Gilpin EA, et al. Which adolescent experimenters progress to established smoking in the United States? *Am J Prev Med.* 1997;13:385–391.

29. Hamilton WL, Biener L, Rodger CN. Who supports tobacco excise taxes? Factors associated with towns' and individuals' support in Massachusetts. *J Public Health Manag Pract.* 2005;11:333–340.

30. Census of Population and Housing, 2000: Summary Tape File 3A. Washington, DC: US Bureau of the Census; 2002.

31. Raudenbusch SW, Bryk AS. *Hierarchical Linear Models: Applications and Data Analysis Methods.* 2nd ed. Thousand Oaks, CA: Sage Publications; 2002.

32. Hedeker D, Gibbons RD, Flay BR. Random-effects regression models for clustered data with an example from smoking prevention research. *J Consult Clin Psychol.* 1994;62:757–765.

33. Szabo E, White V, Hayman J. Can home smoking restrictions influence adolescents' smoking behaviors if their parents and friends smoke? *Addict Behav.* 2006;31:2298–2303.

CLINICAL UPDATE

Update in Pain Medicine

Daniel P. Alford, MD, MPH[1], Jane Liebschutz, MD, MPH[1,2], Ian A. Chen, MD, MPH[3], Christina Nicolaidis, MD, MPH[4], Mukta Panda, MD[5], Karina M. Berg, MD, MS[6], Jennifer Gibson, MD[7], Michael Picchioni, MD[8], and Matthew J. Bair, MD, MS[9]

[1]Boston University School of Medicine, Clinical Addiction Research and Education Unit, Section of General Internal Medicine, Boston Medical Center, Boston, MA, USA; [2]Boston University School of Public Health, Boston, MA, USA; [3]Eastern Virginia Medical School, Norfolk, VA, USA; [4]Oregon Health and Science University, Portland, OR, USA; [5]University of Tennessee, Chattanooga, TN, USA; [6]Albert Einstein College of Medicine, Montefiore Medical Center, Bronx, NY, USA; [7]Legacy Portland Hospitals, Portland, OR, USA; [8]Tufts University School of Medicine, Baystate Medical Center, Springfield, MA, USA; [9]Roudebush VA Center of Excellence for Implementing Evidence Based Practice, Regenstrief Institute, and Indiana University School of Medicine, Indianapolis, IN, USA.

KEY WORDS: primary care; chronic pain; opioids; complementary and alternative medicine.
J Gen Intern Med 23(6):841–5
DOI: 10.1007/s11606-008-0570-8

INTRODUCTION

More than 75 million Americans have chronic or recurrent pain.[1] Pain accounts for 20% of all outpatient visits[2] and more than $100 billion dollars per year in direct (i.e., health care services) and indirect costs (i.e., lost productivity)[3]; analgesics account for 12% of all prescriptions.[4] Chronic pain is a leading cause of work loss, and disability and is a common reason for use of alternative medicine.[5] Our aims were to: review recent pain medicine studies and their key findings and understand how these new findings may impact generalist clinical practice.

We used a systematic search strategy for the period of January 1, 2006 through March 31, 2007 for human subject, English language, peer-reviewed articles that could potentially change generalist care of patients with chronic pain. We searched MEDLINE and PubMed using the medical subject heading (MeSH) terms *pain*, *chronic pain*, and *primary care*. Members of the Society of General Internal Medicine's Pain Medicine Interest Group also suggested other relevant articles. We narrowed the initial list of 314 references to 33. We independently rated the 33 remaining articles using a 5-point Likert scale (1 = poor to 5 = outstanding) on: impact on general internal medicine clinical practice, clinical policy and research, and the quality of the study methods. Based on ratings and consensus deliberations, we chose a subset of 12 articles. We categorized the articles into 5 topic areas: (1) chronic pain and comorbidities; (2) systems approaches to managing chronic pain; (3) opioids and chronic pain; (4) non-pharmacologic approaches to treating chronic pain; and (5) complementary and alternative pain treatments.

This paper derives from the presentation: Update in Pain Medicine at the 30th annual session of SGIM, April 2007 in Toronto, Canada.
Received July 31, 2007
Revised January 3, 2008
Accepted February 22, 2008
Published online March 11, 2008

CHRONIC PAIN AND COMORBIDITIES

Arnow BA, Hunkeler EM, Blasey CM, et al. Comorbid depression, chronic pain, and disability in primary care. *Psychosomatic Medicine.* 2006;68:262–268.

Major depression and chronic pain frequently coexist [6]. However, the strength of their association is unclear, especially in primary care settings. Arnow et al. conducted a large, cross-sectional survey to estimate the prevalence and strength of association between major depressive disorder (MDD) and chronic pain, and the "clinical burden" (i.e., decrements in health-related quality of life, increased somatic symptoms, and additional mental health illness) associated with these conditions individually and in combination. Participants were recruited from 31 internal medicine and family practice clinics within Kaiser Permanente Health Maintenance Organization (HMO) of Northern California. Eligible patients (n=10,710), randomly selected within 1 week of their clinic visit, were mailed a survey. Data from 5,808 respondents (54%) were analyzed. Assessments included psychiatric disorders[7] (depression, anxiety, and alcohol abuse or dependence), somatic symptom severity, health-related quality of life (HRQL), pain-related disability, and chronic pain. Chronic pain was dichotomized as "non-disabling" and "disabling."

Seven percent of respondents met criteria for MDD and 45% experienced chronic pain (28% had disabling pain). Among those with MDD, a significantly higher proportion reported chronic pain compared to those without MDD (66% vs. 43%). Coexisting MDD and chronic pain were associated with poorer HRQL, greater somatic symptom severity, and higher prevalence of panic disorder. The prevalence of alcohol abuse or dependence was two times higher in those with MDD compared to those without MDD. Anxiety disorders were six times more prevalent in those with MDD versus those without regardless of pain presence or disability level.

In summary, chronic pain is especially common among those with MDD. Additionally, the combination of MDD and chronic pain are associated with greater decrements in HRQL, more somatic preoccupation, and more frequent psychiatric comorbidity than MDD alone. The study was limited by a 54% response rate and restricted to patients with a recent clinic visit within an HMO. However, these findings strongly suggest that attention to the assessment and treatment of depression and chronic pain concurrently may be necessary to reduce the clinical burden associated with these conditions.

SYSTEMS APPROACHES TO MANAGING CHRONIC PAIN

Wiedemer NL, Harden PS, Arndt IO, Gallagher RM. The opioid renewal clinic: a primary care, managed approach to opioid therapy in chronic pain patients at risk for substance abuse. Pain Medicine 2007;8:573–84.

Despite limited training in pain medicine, primary care providers (PCPs) manage the bulk of patients with chronic pain. Opioid analgesics are gaining wider acceptance by PCPs, but are controversial for "at risk" patients with a history of substance use disorder or aberrant behavior.

Wiedemer et al. conducted a naturalistic prospective outcome study to measure the impact of a structured opioid renewal program for at risk patients with chronic pain requiring opioids. The study was conducted at the primary care clinic at the Philadelphia Veterans Affairs Medical Center. The intervention involved regular assessments and monitoring by a clinical pharmacist and a nurse practitioner that worked as a liaison between primary care and a multidisciplinary pain team. In addition, PCPs were trained in the use of opioid agreements and random drug testing. Outcomes included providers' use of and patients' adherence to opioid agreements and drug testing, provider satisfaction, and pharmacy costs.

Of 335 patients referred to the program, 171 (51%) had documented aberrant behaviors (e.g., positive drug test), and 164 (49%) had a history of substance use disorder. In those with documented aberrant behaviors, 38% self-discharged from the program, 13% were referred for addiction treatment, and 4% were weaned off for consistently negative urine for prescribed opioids. Of the patients with a history of substance use disorder but no documented aberrant behaviors at the outset, all were adherent to the program. PCP's use of opioid treatments agreements increased fourfold and random drug testing increased substantially. PCPs expressed high levels of satisfaction with the program and significant pharmacy savings were shown.

The study was limited by lack of a comparison group. However, it demonstrated that a nurse practitioner/clinical pharmacist-run clinic, supported by a multi-specialty pain team, can facilitate the use of widely accepted tools such as opioid treatment agreements and urine drug screens by primary care providers in managing opioids in at risk chronic pain patients.

Ahles T, Wasson J, Seville J et al. A controlled trial of methods for managing pain in primary care patients with or without co-occurring psychosocial problems. *Ann Fam Med* 2006:4;341–350.

Behavioral treatments proven to aid pain outcomes include self-management, cognitive–behavioral therapy, and problem-solving therapy.[8,9]. PCPs are not trained to deliver these effective behavioral treatments especially in patients with psychosocial problems. Ahles et al. tested a "stepped" behavioral approach employing individualized self-management skills and problem-solving therapy for pain management in primary care.

The study was a randomized controlled trial in a rural practice-based research network for patients with at least moderate pain lasting for >1 month. Randomization was stratified by the presence or absence of psychosocial problems (self-reported impairment: emotional problems, social activities, social support, sexual problems, substance abuse or household violence). Patients without psychosocial problems (n=693) were randomized to self-management information delivered during a computer feedback session or usual care. The computer-generated feedback targeted both patients and their physicians and provided information from a self-care educational booklet. Patients with psychosocial problems (n=644) were randomized to three arms: computer feedback session alone, computer feedback plus nurse-educator-delivered intervention by phone over 6 months, or usual care. The nurse-educator intervention included: (1) assessment of pain, psychosocial problems, and management preferences, (2) self-management strategies, (3) problem-solving approach, and (4) feedback to the PCP.

The main outcomes included Medical Outcomes Study 36-Item Short-Form (SF-36)[10] domain scores, functional interference, and health care utilization. The participants had a mean age of mid-40s, and most were white, female, married, educated, and employed. The computer-generated feedback did not improve any outcomes in patients at 12-month follow-up. Compared to the usual care control group, the computer feedback plus nurse-educator intervention showed statistically significant improvements (p<.05) for all subscales of the SF-36 except for physical function and social function, with clinically relevant score increases of 6 to 12.5 at 6 months. At 12 months, subscales showed clinically relevant score increases of 5 to 13.9, although only changes in vitality and role emotional remained significant. At 6 and 12 months, functional interference scores were significantly improved. There, were no differences in utilization, although the study was underpowered to show small differences in the overall low utilization. A telephone-based nurse-educator intervention may be a useful treatment program for patients with chronic pain and psychosocial problems.

Mularski RA, White-Chu F, Overbay D, et al. Measuring pain as the 5th vital sign does not improve quality of pain management. *J Gen Intern Med.* 2006; 21:607–612.

The Veterans Health Administration launched the "Pain as the 5th Vital Sign" (P5VS) initiative in 1999 to improve pain management for veterans. The P5VS initiative required the assessment of pain intensity (0 to 10) at all clinical encounters. Mularski et al. sought to measure the initiative's impact on the quality of pain management in a general internal medicine clinic.

Medical records of 300 randomly selected patient visits were reviewed before and after implementation of the P5VS initiative. Seven process indicators were assessed to measure the quality of pain management. A subgroup analysis of patients reporting "substantial pain," defined as a pain score of 4 or greater, was also performed.

Even though pain intensity was documented more frequently (82% vs. 31%) after the initiative, the quality of pain care was unchanged after implementation. There were no significant differences among the process indicators of provider assessment, pain exam, orders to assess pain, new analgesic prescribed, change in existing analgesics, other pain treatment, or follow-up plans. Patients (n=79) who reported substantial pain often did not receive recommended care: 22% had no pain processes documented in the medical record, 27% had no further assessment, and 52% received no new pain therapy at that visit.

The study suggests that a simple pain intensity score assessment is insufficient to improve the evaluation and treatment of patients' pain.

Sullivan MD, Leigh J, Gaster B. Training internists in shared decision making about chronic opioid treatment for noncancer pain. *J Gen Intern Med.* 2006;21:360–62.

Long-term opioid use for chronic pain is a controversial issue in primary care. Shared decision-making models have been shown to improve patient-centered care[11] and may improve care for pain. This study tested a shared decision-making model for opioid treatment of chronic pain in primary care and whether the model improved physician satisfaction and quality of care for patients with chronic pain.

Sullivan et al. conducted a randomized controlled trial of internal medicine residents (n=38) and attendings (n=7). Study participants were randomized to two 1-hour training sessions versus written educational materials of opioid management. Training sessions focused on applying the shared decision-making model with patients when discussing treatment goals, non-medication pain treatments, prescription of methadone as the long-acting opioid of choice, and the role of depression and its treatment in chronic pain. Three-month outcomes included provider satisfaction, degree of patient-centered treatment (doctor receptiveness, patient involvement, affective content of the relationship and information giving), and pain management practices (i.e., methadone prescribing, setting functional goals, opioid treatment agreements).

Compared to the control arm, participants in the shared decision-making model arm reported improvements in: overall satisfaction including relationship quality and appropriate use of time. The intervention group was significantly more likely to give patients information to assist with decision making, prescribe methadone, set functional goals, and complete opioid treatment agreements.

Training in a shared decision-making model improved attitudes and behaviors related to opioid treatment of chronic pain. Pain severity and function were not assessed, and self-reported behaviors were not confirmed. Study physicians' sense of collaboration and satisfaction in treating chronic pain patients was improved by a shared decision-making model.

OPIOIDS AND CHRONIC PAIN

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: 116–127.

Back pain, the second leading symptom seen by US physicians, substantially impacts HRQL.[12] While opioids effectively treat acute pain, benefits of long-term use are unclear. Martell et al. systematically reviewed English-language studies from 1966 to March, 2005 to determine: (1) the prevalence of opioid treatment, (2) efficacy of opioids, and (3) the prevalence of substance use disorders among patients receiving opioids for chronic back pain.

In 11 studies, opioid prescribing varied by treatment setting: 11% to 66% in specialty settings and 3% to 31% in primary care. Opioid prescribing was more common in patients reporting higher disability, worse suffering, and poorer functioning but not necessarily higher pain levels. Of the 15 studies evaluating efficacy, 6 compared an opioid with a non-opioid or placebo, and 9 compared different opioids. All studies had heterogeneous study designs, none lasted more than 4 months, and 11 were industry sponsored. Of the studies comparing an opioid with a non-opioid or placebo, 4 found a non-significant pain reduction with opioids. The 5 most rigorously conducted studies comparing different opioids found a non-significant pain reduction from baseline. Across 5 studies measuring opioid misuse, the prevalence was 5% to 24%, and in 9 studies that assessed current and lifetime substance use disorder, the prevalence was 3% to 43% and 36% to 54%, respectively.

The efficacy of long-term opioids for chronic back pain remains unclear, while evidence does exist to support this treatment for short term use, that is, less then 4 months. Substance use disorders are common in patients taking opioids for back pain. However, the true prevalence of addiction (preexisting or iatrogenic) is still unknown. Despite common use of opioids for chronic back pain, this systematic review cannot provide evidence of long-term efficacy. In addition, evidence about developing addiction from prescribed opioids is too limited to draw any conclusions.

Olsen Y, Daumit G, Ford D. Opioid prescriptions by US primary care physicians from 1992–2001. *J Pain*, 2006; 7:225–235.

Little is known about PCPs opioid prescribing practices. This study assessed trends and factors associated with opioid prescribing of US PCPs nationwide over a 10-year period.

Olsen et al. analyzed cross-sectional demographic, clinical, and prescription data from the 1992–2001 National Ambulatory Medical Survey. Yearly response rates ranged from 63% to 73%; representing between 1,801 and 2,587 physicians and 20,760–36,875 patient visits. Only visits to PCPs were included. The primary outcome was prevalence of primary care visits (per 1,000) in which an opioid was prescribed. The analysis was adjusted for ethnicity, geography, and insurance status.

The prevalence of visits (per 1,000 visits) during which an opioid was prescribed increased from a low of 41 in 1992–1993 to a peak of 63 in 1998–1999 (p<.0001 for trend) and then dropped to 59 in 2000–2001. Several factors increased the adjusted odds ratio (aOR) of receiving opioids: Medicaid [aOR= 2.09, (1.82–2.4)], Medicare [aOR=2.0, (1.68–2.39)], a visit between 15–35 minutes [aOR=1.16, (1.05–1.27)], and receiving NSAIDs [aOR=2.27, (2.04–2.53)]. Factors which lowered the odds of receiving opioids included: Hispanic race [aOR= 0.67, (0.56–0.81)], other race, i.e., Asian/Native American [aOR=0.68, (0.52–0.90)], participation in a HMO [aOR=0.74, (0.66–0.84)], living in the Northeast [aOR=0.6, (0.510.69)], or the Midwest [aOR=0.75, (0.66–0.85)].

Study limitations included the cross-sectional design and lack of information about disease severity within groups being compared. However, this study demonstrates that substantial variations in opioid prescribing practices exist among PCPs, suggesting differences in the quality of pain management across the United States.

Ives TJ, Chelminski PR, Hammett-Stabler CA et al. Predictors of opioid misuse in patients with chronic pain: a prospective cohort study. *BMC Health Services Research.* 2006; 6:46.

Despite growing public health concerns of opioid misuse and addiction, there is scant information regarding the prevalence and risk of opioid analgesic misuse in clinical populations.[13] This prospective cohort study estimated the prevalence and predictors of opioid misuse in an academic internal medicine practice.

Ives et al. followed 196 primary care patients in a chronic pain management program for 1 year. Patients were referred to the

program with difficult to manage pain or suspicions of opioid misuse. A multidisciplinary team (clinical pharmacist, internist, psychiatrist, nurse, and program assistant) developed a multimodality (pharmacologic and non-pharmacologic) pain management plan for each patient in conjunction with the PCP. The primary outcome was opioid misuse defined as: a negative urine toxicology screen (UTS) for the prescribed medications; or positive UTS for non-prescribed opioids, cocaine, or amphetamines; or evidence of multiple prescriptions from different providers, prescription diversion, or forgery.

Most patients were white, with a mean age of 52 years, and nearly half were women. Patients were predominantly low income; more than half were disabled and 29% had a history of substance abuse disorder. One-year data were available for 96% of the patients. Opioid misuse occurred in 32% of patients. The significant factors associated with opioid misuse on multivariate analyses were age (aOR=0.95, 95% CI 0.90–0.99), prior driving under the influence or drug convictions (aOR=2.58, 1.01–6.59), history of cocaine abuse (aOR=4.30, 1.76–10.4), and history of alcohol abuse (aOR=2.60, 1.12–6.26).

The selective nature of the study population limits the generalizability of the prevalence estimate. However, this study characterizes factors associated with opioid misuse and provides a practical working definition of misuse that emphasizes the role of UTS in monitoring patients on long-term opioid therapy. Access to prior drug- and alcohol-related conviction data may be useful when available.

NON-PHARMACOLOGIC APPROACHES TO TREATING CHRONIC PAIN

Weinstein JN, Tosteson TD, Lurie JD, et al. Surgical vs nonoperative treatment for lumbar disk herniation: the spine patient outcomes research trial (SPORT): a randomized trial. *JAMA* 2006:296(20):2441–2450.

Weinstein JN, Lurie JD, Tosteson TD, et al. Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT): observational cohort study. *JAMA* 2006:296(20):2451–2459.

Lumbar diskectomy is the most common surgery performed for back and radicular leg symptoms in U.S. patients.[14] However, controversy exists regarding its efficacy compared to non-operative care.

The Spine Patient Outcomes Research Trial (SPORT) studies assessed the efficacy of surgery compared to non-operative care for lumbar intervertebral disk herniation. The study involved a concurrent randomized controlled trial (RCT) and an observational cohort study and included 13 interdisciplinary spine clinics in 11 U.S. states. Participants were surgical candidates with imaging-confirmed lumbar disk herniation and signs and symptoms of radiculopathy lasting at least 6 weeks. Patients undergoing operative diskectomy vs. non-operative therapy were compared. Non-operative therapy included physical therapy, education/counseling with home exercise instruction, and NSAIDs, if tolerated. Physicians caring for patients in the non-operative arm were also provided with a list of other therapies and encouraged to individualize treatment. Main outcomes included bodily pain and physical function as measured by the SF-36[10] back-pain-specific physical function as measured by the Oswetry Disability Index[15], sciatica severity, satisfaction, self-reported improvement, and employment status at 3, 6, 12, and 24 months. Participants who refused randomization at baseline were entered into the cohort study.

In the RCT, 232 participants were randomized to surgery, but half (50%) did not undergo surgery. Of the 240 participants randomized to non-operative care, 30% had surgery. Both groups showed significant improvements. The intention to treat analysis favored surgery, but all differences were small and not statistically significant.

In the cohort study, 521 participants chose surgery and 96% underwent the procedure. Of those who chose non-operative care (n=222), 22% eventually underwent surgery. Both groups improved over time, but the surgery arm showed greater improvements on all measures. Improvements were clinically significant, with a 15-point difference between groups on both SF-36[10] scales for bodily pain and physical function at 3 months. Differences narrowed slightly, but persisted for the entire 2-year study period.

Due to the large degree of cross-over in the RCT and potential selection bias in the observational study, conclusions regarding the superiority or equivalence of treatments are not warranted. Because all patients had imaging-confirmed, symptomatic, and persistent disk herniations, not simply low back pain, the results should not be generalized to the broader population of patients with chronic low back pain.

It is unlikely that we will have a clearer answer in the near future, as it will be hard to improve on the methods of this rigorously conducted, well-funded large trial. However, providers may be reassured that most back pain improves—even for patients that meet strict criteria for disc surgery—and that following patient preference may be a reasonable and evidence-based approach.

COMPLEMENTARY AND ALTERNATIVE PAIN TREATMENTS

Brinkhaus B, Witt CM, Jena S, et al. Acupuncture in patients with chronic low back pain: a randomized controlled trial. *Arch Intern Med.* 2006;166:450–457.

Despite the lack of evidence for complementary and alternative medicine (CAM) treatments for pain conditions, one third of U.S. adults with low back pain seek pain relief using CAM, including acupuncture.[5] The Acupuncture Randomized Trial in Low Back Pain, a multi-center trial, tested the efficacy of acupuncture in reducing chronic low back pain.

Eligible patients were aged 40 to 75 years, with greater than 6 months of low back pain of unclear etiology, moderate pain intensity, and receiving only NSAID analgesia. Participants were randomized to acupuncture (n=146), sham acupuncture (n=73), or wait-list control (n=79). The primary outcome was change in pain intensity on visual analog scale at 8 weeks. By week 8, acupuncture significantly decreased pain compared to wait-list control, but not to the sham acupuncture. Results remained similar at 26 and 52 weeks for all outcome measures. Fifteen patients (11%) receiving acupuncture and 12 patients (17%) receiving sham acupuncture (p=.20) reported adverse effects, including hematoma and bleeding.

The authors concluded that acupuncture (including sham acupuncture) was more effective than no acupuncture (wait list) in patients with chronic low back pain. This is one of the

largest and most rigorous trials to investigate the efficacy of acupuncture for low back pain. Study strengths included: assessment of intervention credibility, interventions delivered by qualified and experienced medical acupuncturists, and high follow-up rates.

The lack of difference between acupuncture and sham acupuncture suggests that sham acupuncture may also have specific analgesic effects that need further exploration. Expectation bias (active treatment vs. wait list) and placebo (active vs. sham were similar) effects could confound outcomes. Future head-to-head trials comparing acupuncture and other interventions for treating chronic low back pain are needed.

Clegg DO, Reda DJ, Harris CL, et al. Glucosamine, chondroitin sulfate, and the two in combination for painful knee osteoarthritis *NEJM*. 2006; 354(8):795–808.

The dietary supplements glucosamine and chondroitin sulfate have been promoted as safe and effective treatment options for osteoarthritis symptoms. A meta-analysis of studies evaluating these supplements suggested potential benefit, but questioned the quality of included studies.[16] The Glucosamine/Chondroitin Arthritis Intervention Trial (GAIT) was a 24-week randomized double-blind, placebo- and celecoxib-controlled multi-center trial to evaluate the efficacy and safety of glucosamine, chondroitin sulfate, and glucosamine plus chondroitin sulfate in the treatment of painful knee osteoarthritis.

The GAIT study included adult patients who had clinical and radiographic evidence of knee osteoarthritis, had an elevated Western Ontario and McMaster Osteoarthritis Index (WOMAC)[17] pain scores, and were physically functional. The 1,583 eligible patients were randomized to receive daily doses of 1,500 mg glucosamine, 1,200 mg chondroitin sulfate, both glucosamine and chondroitin sulfate, 200 mg of celecoxib, or placebo for 24 weeks. The primary outcome was a reduction in the WOMAC pain scale of 20%.

Overall, glucosamine and chondroitin sulfate were not significantly better than placebo in reducing knee pain by 20%. Compared to placebo (60.1%), the response to glucosamine was 64% (p=.30), to chondroitin sulfate was 65.4% (p=.17), to combined treatment was 66.5% (p=.09), and to celecoxib was 70.1% (p=.008). Subgroup analysis of patients with moderate-to-severe pain demonstrated that combination therapy significantly decreased pain compared to placebo (p=.002). Adverse events were infrequent and mild and evenly distributed among the groups.

The large placebo response and relatively mild degree of pain from osteoarthritis among the participants may have limited the ability to detect a difference in treatment efficacy. While glucosamine and chondroitin sulfate alone or in combination did not show efficacy in the overall study group, combination therapy may have efficacy in patients with more severe symptoms from knee osteoarthritis.

Acknowledgments: *All authors are members of the Society of General Internal Medicine's Pain Medicine Interest Group. The authors also wish to thank Amy Alawad at Boston Medical Center for her administrative support in coordinating the submission of this manuscript.*

Conflicts of Interest: *Dr. Bair has received honoraria from CME Enterprises in an educational program sponsored by Eli Lilly and Company. The other authors reported no potential conflicts of interest.*

Corresponding Author: *Daniel P. Alford, MD, MPH; Clinical Addiction Research and Education Unit, Section of General Internal Medicine, Boston Medical Center, 801 Massachusetts Avenue, 2nd Floor, Boston, MA 02118, USA (e-mail: dan.alford@bmc.org).*

REFERENCE

1. National Center for Health Statistics. Health, United States, 2006, with chartbook on trends in the health of Americans. 68–71. 2006. Hyattsville, MD.
2. **Schappert SM.** National Ambulatory Medical Care Survey: 1989 summary. Vital Health Stat. 1992;13:1–80.
3. **Turk DC, Okifuji A, Kalauokalani D.** Clinical outcome and economic evaluation of multidisciplinary pain centers. In: **Block AR, Kremer EF, Fernandez E**, eds. Handbook of pain syndromes: Biopsychosocial perspectives. Mahwah, NJ: Erlbaum; 1999:77–98.
4. **Turk DC.** Clinical effectiveness and cost-effectiveness of treatments for patients with chronic pain. Clin J Pain. 2002;18:355–65.
5. **Eisenberg DM, Davis RB, Ettner SL, et al.** Trends in alternative medicine use in the United States, 1990–1997: results of a follow-up national survey. JAMA. 1998;280:1569–75.
6. **Bair MJ, Robinson RL, Katon W, Kroenke K.** Depression and pain comorbidity: A literature review. Arch Intern Med. 2003;163:2433–45.
7. **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. 1999;282:1737–44.
8. **Morley S, Eccleston C, Williams A.** Systematic review and meta-analysis of randomized controlled trials of cognitive behaviour therapy and behaviour therapy for chronic pain in adults, excluding headache. Pain. 1999;80:1–13.
9. **van den Hout JH, Vlaeyen JW, Heuts PH, Zijlema JH, Wijnen JA.** Secondary prevention of work-related disability in nonspecific low back pain: does problem-solving therapy help? A randomized clinical trial. Clin J Pain. 2003;19:87–96.
10. **Ware JE Jr.** SF-36 Health Survey. Manual and interpretation guide. Boston, MA: Nimrod Press; 1993.
11. **Charles C, Gafni A, Whelan T.** Shared decision-making in the medical encounter: What does it mean? (or it takes at least two to tango). Soc Sci Med. 1997;44:681–92.
12. **Deyo RA, Weinstein JN.** Low back pain. N Engl J Med. 2001;344:363–70.
13. **Compton WM, Volkow ND.** Abuse of prescription drugs and the risk of addiction. Drug Alcohol Depend. 2006;83(Suppl 1):S4–S7.
14. **Weinstein JN, Lurie JD, Olson PR, Bronner KK, Fisher ES.** United States' trends and regional variations in lumbar spine surgery: 1992–2003. Spine. 2006;31:2707–14.
15. **Daltroy LH, Cats-Baril WL, Katz JN, Fossel AH, Liang MH.** The North American spine society lumbar spine outcome assessment Instrument: Reliability and validity tests. Spine. 1996;21:741–49.
16. **McAlindon TE, LaValley MP, Gulin JP, Felson DT.** Glucosamine and chondroitin for treatment of osteoarthritis: A systematic quality assessment and meta-analysis. JAMA. 2000;283:1469–75.
17. **Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW.** Validation study of WOMAC: A health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol. 1988;15:1833–40.

BMC Medical Education

BioMed Central

Research article

Open Access

A web-based Alcohol Clinical Training (ACT) curriculum: Is in-person faculty development necessary to affect teaching?

Daniel P Alford*[1,2], Jessica M Richardson[1], Sheila E Chapman[1,2], Catherine E Dubé[3], Robert W Schadt[4] and Richard Saitz[1,2,5,6]

Address: [1]Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Department of Medicine, Boston Medical Center, Boston, MA, USA, [2]Boston University School of Medicine, Boston, MA, USA, [3]Department of Community Health, Brown University, Providence, RI, USA, [4]Department of Educational Technology, Boston University School of Public Health, Boston, MA, USA, [5]Youth Alcohol Prevention Center, Boston University School of Public Health, Boston, MA, USA and [6]Department of Epidemiology, Boston University School of Public Health, Boston, MA, USA

Email: Daniel P Alford* - dalford@bu.edu; Jessica M Richardson - jesricha@montefiore.org; Sheila E Chapman - schapman@bu.edu; Catherine E Dubé - Catherine_Dube@Brown.edu; Robert W Schadt - rschadt@bu.edu; Richard Saitz - rsaitz@bu.edu

* Corresponding author

Published: 6 March 2008

BMC Medical Education 2008, **8**:11 doi:10.1186/1472-6920-8-11

Received: 13 June 2007
Accepted: 6 March 2008

This article is available from: http://www.biomedcentral.com/1472-6920/8/11

Abstract

Background: Physicians receive little education about unhealthy alcohol use and as a result patients often do not receive efficacious interventions. The objective of this study is to evaluate whether a free web-based alcohol curriculum would be used by physician educators and whether in-person faculty development would increase its use, confidence in teaching and teaching itself.

Methods: Subjects were physician educators who applied to attend a workshop on the use of a web-based curriculum about alcohol screening and brief intervention and cross-cultural efficacy. All physicians were provided the curriculum web address. Intervention subjects attended a 3-hour workshop including demonstration of the website, modeling of teaching, and development of a plan for using the curriculum. All subjects completed a survey prior to and 3 months after the workshop.

Results: Of 20 intervention and 13 control subjects, 19 (95%) and 10 (77%), respectively, completed follow-up. Compared to controls, intervention subjects had greater increases in confidence in teaching alcohol screening, and in the frequency of two teaching practices – teaching about screening and eliciting patient health beliefs. Teaching confidence and teaching practices improved significantly in 9 of 10 comparisons for intervention, and in 0 comparisons for control subjects. At follow-up 79% of intervention but only 50% of control subjects reported using any part of the curriculum ($p = 0.20$).

Conclusion: In-person training for physician educators on the use of a web-based alcohol curriculum can increase teaching confidence and practices. Although the web is frequently used for disemination, in-person training may be preferable to effect widespread teaching of clinical skills like alcohol screening and brief intervention.

Background

Practice guidelines of leading professional societies recommend alcohol screening and behavioral counseling interventions in primary care settings [1-3]. Valid, brief,

practical screening tools exist for the detection of unhealthy alcohol use in primary care settings[4], and brief interventions by physicians can reduce drinking and improve health outcomes when delivered to primary care patients with unhealthy alcohol use [5-8]. However, unhealthy alcohol use in primary care is often unrecognized and untreated, as reported in studies performed well after research demonstrating efficacy and national guidelines were published [9-14]. Although physicians recognize their responsibility in identifying and addressing alcohol problems [15], it often does not occur using effective patient-centered techniques [16]. Physician avoidance of and discomfort with brief alcohol counseling have been identified as important barriers [17].

Physician education can improve screening and brief intervention skills resulting in decreased patient drinking [7,18-22]. Some education and training programs aimed at improving physician attitudes and clinical practice around substance abuse issues have been effective [23-32]. However, despite the existence of numerous curricula [33], they are not being widely used [34]. Only half of internal medicine residency training programs have training on initial diagnosis and management of substance use disorders [34].

Web-based training can be an innovative and efficient way to connect with many individuals, while allowing learning at a convenient time for the learner. Adult learning principles [35] suggest that physicians' use of information sources outside the local sphere, such as journals, conferences, and the Internet, are essential to the enhancement and acceleration of information diffusion throughout the medical community. Although journals and books are the most common mechanism by which research findings are disseminated, they are not always read by practicing physicians [36]. The Internet can provide flexible, adaptable, tailored and sustainable access to current information [37-41] allowing for self-directed and individualized learning. Physicians have come to rely on the Internet for accessing clinical information [42] and for continuing medical education (CME) [43]. Internet-based CME has been shown to improve physician knowledge and change physician behavior [44-48]. However, little data is available on the use of web-based curricular materials by physician educators or the effectiveness of faculty development programs aimed at increasing physician use of the Internet curriculum resources. Further, although the train-the-trainer model is an efficient and widely accepted mechanism of curriculum dissemination, it is not known whether and to what degree such efforts enhance physician use of web-based curriculum tools.

To enhance dissemination of alcohol skills training to physicians, we developed an easily transportable curriculum that meets the general requirements of successful web-based courses [49] and adult learning theories [35] that could be actively distributed, easily integrated into existing curricula and used by internal medicine faculty educators. In this study, we tested whether in-person faculty development training is associated with a) use of a free web-based Alcohol Clinical Training (ACT) curriculum among physician educators, b) increased alcohol-related teaching confidence, and c) increased specific alcohol-related teaching practices.

Methods

The ACT curriculum

The Alcohol Clinical Training (ACT) curriculum is a federally funded, web-based curriculum created specifically for general internist educators to teach improved clinical and communication skills (screening, assessment and brief intervention) important in addressing unhealthy alcohol use in primary care settings. The ACT curriculum is based on the U.S. National Institute on Alcohol Abuse and Alcoholism (NIAAA) *Helping Patients Who Drink Too Much: A Clinician's Guide* [50]. With a special focus on health disparities, curricular topics include the spectrum of alcohol use, selected health consequences of alcohol use, epidemiology of unhealthy alcohol use, alcohol problems frequently missed, effects of physician culture on doctor/patient communication, screening, and brief intervention. The ACT curriculum was developed by and for general internists and is designed for teaching faculty, residents and medical students in a variety of teaching settings including small group conferences and large group didactic sessions. It consists of PowerPoint slides with case-based video vignettes, as well as speaker notes and audio, and learner evaluation materials. The curriculum is designed to be flexible and modifiable (i.e. slide content can be changed and videos are available as streaming or downloadable files) and can be taught using all the components together in a 3 hour workshop or by using various components separately in 45 minute sessions (i.e. preclinic conference or attending rounds).

Pilot studies

Pilot testing was conducted to fine tune the Alcohol Clinical Training (ACT) curriculum based on input from learners in real practice settings caring for diverse (economically and culturally) patient populations. Pilot testing was performed with 3 types of physicians including residents in internal medicine, practicing community clinicians, and faculty physician educators.

Study design

In this controlled educational study, we analyzed baseline and 3-month follow-up survey data collected from applicants to a satellite workshop conducted at an American College of Physicians (ACP) national meeting. The study

was approved by the Institutional Review Board at Boston Medical Center.

Subjects were physician educators in the U.S. who applied to attend a workshop on the use of the web-based ACT curriculum. The workshop was advertised on the ACP website and newspaper and via electronic mail to members of several medical professional organizations (e.g. Society of General Internal Medicine, Association of Program Directors in Internal Medicine). The workshop was limited to 20 participants to facilitate an interactive format and because workshop space and resources were limited. Due to this limitation, the first 20 eligible applicants were invited to attend the workshop (intervention subjects), while all eligible applicants who applied after the workshop filled were asked to enroll as control subjects (to complete assessments and have access to the curriculum website but not attend the workshop). The control group was limited by the number of additional applications received, beyond the 20 accepted for the workshop. When intervention subjects attended the workshop, control subjects were sent a letter including a description of, and web address for, the online ACT curriculum [51]. Upon completion of the study, intervention subjects received reimbursement up to $500 for travel costs or a $500 honorarium for completing the baseline assessment, attending the workshop and completing the 3-month follow-up survey; control subjects received $100 for completing the baseline assessment and 3-month follow-up survey. All subjects provided informed consent.

Faculty development workshop

The workshop consisted of 3 hours of in-person, interactive teaching on the effective use of the ACT curriculum, including demonstration of navigating and using website materials, modeling of teaching by expert faculty, and creating an individual action plan: a teaching project focused on using the ACT curriculum within 2 months of the workshop. Participants were required to develop an action plan objective, and to identify the target audience, setting, available resources, potential barriers, and plan for evaluation. Attendees received continuing medical education credits from Boston University.

Assessments

Both intervention and control subjects completed baseline surveys with their applications to attend the workshop. Follow-up surveys were mailed to all subjects 3 months after the workshop. Because up to 5 months separated completion of the baseline survey and workshop attendance, intervention subjects repeated the baseline survey directly preceding the workshop to assess whether baseline results changed (e.g. due to secular trend or in response to being selected to attend the workshop) (Figure 1: Participation Summary).

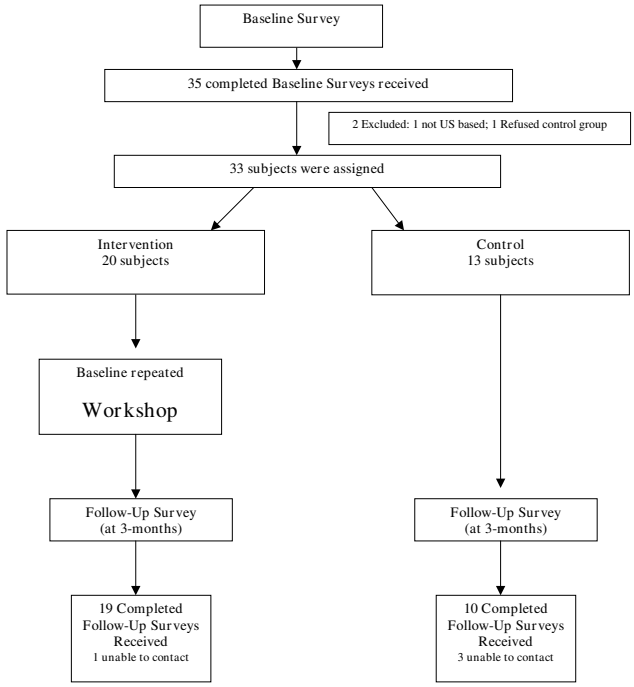


Figure 1
Participation Summary.

Baseline surveys included questions on respondent characteristics such as: demographics (gender, race, ethnicity, first language [English: yes/no], age, number of fluent languages other than English), residency completion year, primary teaching settings and expertise in the diagnosis and management of alcohol problems (yes/no) with any affirmative response to "Do you have expertise in the diagnosis and management of alcohol problems through: American Society of Addiction Medicine (ASAM) certification, past faculty fellowship(s), practice in an addictions specialty setting or other specified being counted as substance abuse "expertise"." Additionally, baseline surveys assessed the settings in which subjects had taught about alcohol problems (i.e., resident conferences/seminars, medical student courses or conferences, continuing medical education (CME) courses, grand rounds, morning report, inpatient attending rounds, teaching while providing clinical care, other). Note that although the term "unhealthy alcohol use" better encompasses the spectrum of use of clinical interest than the term "alcohol problems," we use the latter in describing our methods and results because it was the term in use at the time of the study (consistent with the contemporaneous NIAAA guideline) [52].

Based on work by D'Onofrio and colleagues [21], the subjects were asked at both baseline and follow-up to rate, on

a 5-point Likert Scale, their teaching confidence (from "not at all" to "very") and specific teaching practices (from "rarely" to "always") in the following 5 domains: alcohol screening, assessment of readiness to change, counseling about alcohol problems, eliciting patient health beliefs, and assuring patients that they are understood [21] Follow-up surveys assessed which components (slides, notes, audio, any, none) and in which settings the ACT curriculum was used in the prior 3-months. Intervention subjects were provided a copy of their action plan and asked how much of it had been completed (none, some, or all).

Outcomes

The primary outcomes are baseline to follow-up change in self-reported teaching confidence (5 domains) and specific teaching practices (5 domains). Secondary outcomes included curriculum use, type of teaching settings, and frequency of alcohol-related advice sought. Degree of action plan completion was an outcome for intervention subjects only.

Statistical analyses

All data was analyzed with SAS/STAT software, Version 8.2 [53]. Initial analyses consisted of descriptive statistics (means, standard deviations, medians, interquartile ranges, and proportions). Comparisons were performed with 2 sample t-tests for continuous variables and chi square tests for categorical variables. Reported p-values are two-tailed, and a p-value less than 0.05 was considered statistically significant.

For the primary outcomes, we compared mean change from baseline to follow-up (calculated as follow-up score minus baseline score) in the 5 domains of teaching confidence and 5 domains of specific teaching practices both between and within groups. We also compared between-group differences in secondary outcomes. Lastly, for intervention subjects, we compared responses on the baseline application to those from the pre-workshop repeat baseline surveys.

Results

Of the 35 physicians who completed baseline surveys, 1 was not U.S.-based and 1 refused to participate in the control group; thus 33 were enrolled (Figure 1). Of 20 intervention and 13 control subjects, 19 (95%) and 10 (77%), respectively, completed follow-up. There were no statistically significant differences in baseline characteristics by group (Table 1) including self-reported counseling and teaching others to counsel patients with alcohol problems regarding their alcohol use. One subject was missing all baseline data, with the exception of gender.

Of the 5 domains of teaching confidence and the 5 domains of specific teaching practices evaluated, compared to controls, intervention subjects increased significantly more in their confidence in teaching alcohol screening (mean change, intervention + 1.24 vs. control + 0.11, p = 0.006) and in the frequency of teaching about alcohol screening (mean change, intervention +0.56 vs. control -0.56, p = 0.02) (Table 2). Intervention subjects also increased significantly more than controls in the frequency of teaching learners to elicit patient health beliefs (mean change, intervention +0.81 vs. control -0.33, p = 0.03). Within group changes from baseline to follow-up in teaching confidence and frequency of specific teaching practices were significant for 9 of the 10 comparisons in intervention subjects, and 0 of the 10 comparisons in control subjects. The intervention subjects' pre-workshop repeat baseline surveys significantly increased compared with the baseline survey in only 1 domain – confidence in teaching to assure patients that they're understood, which did not fully explain the difference between baseline and follow-up scores (baseline to preworkshop repeat baseline mean change +0.51 vs. baseline to follow-up mean change +1.47).

Table 1: Characteristics of the 33 enrolled physician educators

	Intervention Group (N = 20)†	Control Group (N = 13)	p-value
Male (%)	79	62	0.43
Race (%)			0.85
Asian	37	31	
Black/African American	11	8	
White	37	54	
Other	16	8	
Hispanic (%)	5	8	1.00
English First Language (%)	58	54	1.00
Has Substance Abuse Expertise (%)	50	54	1.00
Mean Age	41	45	0.14
Mean # Fluent Languages	2	1	0.37
Mean # Years Since Residency	10	11	0.56

†1 subject missing on all characteristics except gender

Table 2: Baseline to follow-up change in 5 domains of teaching confidence and specific teaching practices

	Intervention (N = 18)†	Control (N = 9)†	Between- group p-value
Teaching confidence§			
Alcohol screening	+ 1.24**	+ 0.11	**0.006**
Assessment of readiness to change	+ 1.00**	+ 0.11	0.06
Counseling about alcohol problems	+ 1.18**	+ 0.44	0.12
Eliciting patient health beliefs	+ 1.29**	+ 0.67	0.23
Assuring patients that they are understood	+ 1.47**	+ 0.56	0.07
Specific teaching practice frequency¶			
Alcohol screening	+ 0.56*	- 0.56	**0.02**
Assessment of readiness to change	+ 0.44	- 0.44	0.09
Counseling about alcohol problems	+ 0.67*	- 0.22	0.08
Eliciting patient health beliefs	+ 0.81**	- 0.33	**0.03**
Assuring patients that they are understood	+ 0.94*	+ 0.11	0.18

*p < .05; **p < .01; in within-group comparisons of baseline to follow-up change
†Baseline data were missing for one subject with follow-up data in each group (1 of 19 in the intervention group and 1 of 10 in the control group)
§5-point Likert scale, where 1 = Not at all Confident and 5 = Very Confident
¶5-point Likert scale, where 1 = Rarely and 5 = Always

At follow-up, there was more curriculum use among intervention subjects than control subjects, though the difference was not significant (79% vs. 50%, p = 0.20) (Table 3). The most commonly used component of the curriculum was the slides, whereas use of the audio component was nearly nonexistent.

Although not intended for this purpose, the curriculum was used for self-learning by the majority of subjects (71%) with no difference between intervention and control groups (Table 4). For teaching, the curriculum was used in a variety of settings, the most common of which were while providing clinical care (61%) and resident teaching conferences (43%).

Among intervention subjects, 84% (16/19) completed at least part of their action plan including 8 participants who completed their entire action plan. Two of the three intervention subjects who did not complete any of their action plan also did not use any part of the curriculum.

Discussion

In-person training for physician educators on the use of a web-based Alcohol Clinical Training (ACT) curriculum is associated with increases in confidence in teaching about alcohol screening and specific teaching practices – more frequent teaching about alcohol screening and eliciting patient health beliefs. Given the small sample size, non-significant increases are also noteworthy, including the increases associated with in-person training in confidence in teaching about assessment of readiness to change and assuring patients that they are understood, and more frequent teaching about assessment of readiness to change and counseling about alcohol problems. Also notable are the within-group findings demonstrating that intervention group confidence and teaching frequency increased significantly in 9 of 10 comparisons, which were not a result of improvements prior to the workshop. In comparison, the control group never improved significantly, and in fact, worsened in some cases. These findings suggest that in-person training of, and not only access to, this web-based curriculum can lead to improvements in alcohol-related teaching confidence and practice.

Table 3: Proportion with curriculum use at follow-up

	Intervention Group (N = 19) N (%)	Control Group (N = 10) N (%)	p-value
Any curriculum use	15 (79)	5 (50)	0.20
Slide Use	11 (58)	4 (40)	0.17
Notes Use	7 (37)	2 (20)	0.26
Audio Use	0 (0)	1 (10)	0.39
Video Use	3 (16)	1 (10)	1.00

Table 4: ACT curriculum teaching settings at follow-up

	Intervention Group (N = 18†) N (%)	Control Group (N = 10) N (%)	p-value
For my own learning	14 (78)	6 (60)	0.42
Resident teaching conferences	9 (50)	3 (30)	0.43
Medical student teaching conferences	7 (39)	1 (10)	0.19
Continuing Medical Education courses	0 (0)	0 (0)	N/A
Grand rounds	2 (11)	0 (0)	0.52
Morning report	5 (28)	1 (10)	0.37
Inpatient attending rounds	9 (50)	1 (10)	0.04*
Teaching while providing clinical care (e.g., precepting)	12 (67)	5 (50)	0.44
Other	2 (11)	0 (0)	0.52

*$p < .05$
† Data missing for 1 subject

Although not significant, a greater proportion of educators who had in-person training used the curriculum. Even among those without in-person training, a substantial proportion of subjects reported using the web-based curriculum. Although physicians used the curriculum, none used the audio portion and very few used the video portion. While the curriculum was used for its intended purpose, as an education tool, surprisingly, it was most commonly used for self-study, even by educators self-selected as having an interest in teaching about alcohol.

Curricular topics for generalist physician educators are expanding in number and scope while residency duration remains the same. Considering the ineffectiveness of medical residency programs in training for alcohol screening and management [34,54], target audiences for this curriculum include both the physicians who will use the curriculum to train others, and physicians being trained. The intent of developing and making this alcohol education curriculum available is to provide faculty with a variety of educational materials (i.e. video, slides) that they can take "off the shelf," modify if desired and use in a variety of settings.

The ACT curriculum was developed by and for the same group – general internists. This approach is in keeping with the emphasis on specialty specific teaching in physicians' (adult) learning principles. This approach uses the internist teacher as a role model with credibility specifically applicable to the learner's specialty [55].

Diffusion of alcohol skills training is enhanced or impeded by fundamental characteristics of the training mechanism, such as its complexity and accessibility [56]. Many previously created alcohol curricula for physicians are less easy to access, less tailored to their audiences, and less focused. In the 1990s, NIAAA developed two multimodule curricula, which include materials available for purchase on diskette [57-59]. The Project ADEPT (*Alcohol and Drug Education for Physician Training*) curriculum is a comprehensive substance abuse curriculum for primary care physicians, which includes 7 modules, each with approximately 300 pages of instructional material. Several other non-web-based curricula are similarly lengthy, ranging from 9-hour sessions to 4-day workshops, and often address substance abuse in general, rather than being alcohol-specific, despite the fact that guidelines recommend universal screening for alcohol, but not other drugs. Many require payment or are no longer available. On the other hand, web-based materials are more easily available, and evaluations of web-based physician education have shown significant changes in both non-behavioral measures (e.g. knowledge, attitude, confidence and satisfaction) [44,45,60] as well as behavioral changes [61,62] that impact patient care [44]. With the increasing number of physicians using the Web for continuing medical education [42], it is not surprising that a number of organizations are making curricula available on the web. The NIAAA's *Helping Patients Who Drink Too Much: A Clinician's Guide*, upon which the ACT curriculum is based, is freely available on the Web and in print [50], however, it does not include the audio, video or evaluation features offered with ACT. Project Cork, Clinical Tools, Inc., the University of Florida Division of Addiction Medicine, and likely others, also provide Internet-based alcohol curricula, some of which are focused on screening and brief intervention and are geared toward physicians in general [63,64] while unlike any others, the ACT curriculum is specifically tailored to internist educators.

Several important limitations of this study evaluating the ACT curriculum should be considered. The small sample size makes it difficult to identify differences between groups and caution must be exercised when interpreting the results of non-significant findings. However, some results did reach significance. Second, the nonrandomized nature of the study could have led to confounding. For example, because enrollment in the intervention group was based on early workshop application, intervention subjects might have been more highly motivated

than control subjects. However, we believe individuals in both groups were highly motivated since they all applied to attend a workshop that involved a considerable time commitment. Further, selection bias could lead to difficulty generalizing these results from this group of physician educators volunteered to travel to attend a course to a representative sample of physician educators. Nevertheless, our original intention was to study physicians with an interest in alcohol use and not to generalize beyond that population. Further, it is possible that the workshop learners considered the workshop instructors to be opinion leaders or field experts. Relatedly, subjects' specific teaching practices were not directly measured; as such it is possible that some of the findings may be attributed to social desirability bias, reporting favorable behaviors to researchers evaluating the course they attended.

Despite these limitations, these study results suggest that posting a Web-based curriculum tailored for internist educators can lead to its use, and to improvements in teaching confidence and frequency of teaching practices that are further improved when the curriculum is demonstrated in person. Furthermore, although intended for use by educators to train others, such a curriculum can be used for self-study. More sophisticated enhancements, such as audio and video components, might require more substantial faculty development efforts, and additional research is needed on both how to better disseminate these curricula, and on practice and patient-level outcomes. Nonetheless, this educational tool has the potential, perhaps in conjunction with other efforts [65,66], to improve clinical practice in an area recognized as needing substantial improvement [9].

Conclusion

Leading professional societies recommend that alcohol screening and behavioral counseling interventions be implemented in primary care settings. But physician education to support this implementation has not been effectively or widely disseminated. This study demonstrates that a free web-based alcohol clinical training curriculum will be used by physician educators and that in-person training on the use of the curriculum can further increase teaching confidence and practices.

Competing interests

The author(s) declare that they have no competing interests.

Authors' contributions

DPA and RS led the creation of the study concept and design. All authors made substantial contributions to acquisition, analysis and interpretation of data and were involved in drafting and revising the manuscript for important intellectual content. All authors have read and approved the manuscript for publication.

Acknowledgements

The authors appreciate the data management assistance of Erika Edwards, MPH, and Michael Winter, MPH, at the Data Coordinating Center, Boston University School of Public Health, Boston, MA, USA. Support from this study came from the US National Institute of Alcohol Abuse and Alcoholism (R25 AA13822). Preliminary results were presented at the annual national meetings of the Society of General Internal Medicine (SGIM), April 26–29, 2006 in Los Angeles, CA, USA, the Research Society on Alcoholism, June 24–28, 2006 in Baltimore, MD, USA, and the Association for Medical Education and Research in Substance Abuse (AMERSA), November 2–4, 2006 in Washington, DC, USA.

References

1. **United States Preventive Services Task Force and Screening and behavioral counseling interventions in primary care to reduce alcohol misuse. Recommendation Statement** [http://www.ahrq.gov/clinic/3rduspstf/alcohol/alcomisrs.htm]
2. United States Preventive Services Task Force: *Guide to Clinical Preventive Services* 2nd edition. Washington, DC: Office of Disease Prevention and Health Promotion; 1996.
3. Anderson P, Gual A, Colom J: **Alcohol and Primary Health Care: Clinical Guidelines on Identification and Brief Interventions.** Barcelona, Spain, Department of Health of the Government of Catalonia; 2005.
4. Saitz R: **Unhealthy alcohol use.** *N Engl J Med* 2005, **352:**596-607.
5. Fleming MF, Barry KL, Manwell LB, Johnson K, London R: **Brief physician advice for problem alcohol drinkers: a randomized controlled trial in community-based primary care practices.** *JAMA* 1997, **277:**1039-45.
6. Kristenson H, Ohlin H, Hulten-Nosslin MB, Trell E, Hood B: **Identification and intervention of heavy drinking in middle-aged men: results and follow-up of 24–60 months of long-term study with randomized controls.** *Alcohol Clin Exp Res* 1983, **7:**203-9.
7. Ockene JK, Adams A, Hurley TG, Wheeler EV, Hebert JR: **Brief physician- and nurse practitioner-delivered counseling for high-risk drinkers: does it work?** *Arch Intern Med* 1999, **159:**2198-205.
8. Allen JP, Maisto SA, Connors GJ: **Self-report screening tests for alcohol problems in primary care.** *Arch Intern Med* 1995, **155:**1726-30.
9. McGlynn EA, Asch SM, Adams J, Keesey J, Hicks J, DeCristofaro A, Kerr EA: **The quality of health care delivered to adults in the United States.** *N Engl J Med* 2003, **348:**2635-45.
10. The National Center on Addiction and Substance Abuse at Columbia University: *Missed Opportunity: National Survey of Primary Care Physicians and Patients on Substance Abuse. Conducted by the Survey Research Laboratory, University of Illinois at Chicago* New York, NY: Josiah Macy, Jr. Foundation; 2000.
11. Buchsbaum DG, Buchanan RG, Poses RM, Schnoll SH, Lawton MJ: **Physician detection of drinking problems in patients attending a general medicine practice.** *J Gen Intern Med* 1992, **7:**517-21.
12. Coulehan JL, Zettler-Segal M, Block M, McClelland M, Schulberg HC: **Recognition of alcoholism and substance abuse in primary care patients.** *Arch Intern Med* 1987, **147:**349-52.
13. Rydon P, Redman S, Sanson-Fisher RW, Reid AL: **Detection of alcohol-related problems in general practice.** *J Stud Alcohol* 1992, **53:**197-202.
14. Saitz R, Mulvey KP, Plough A, Samet JH: **Physician unawareness of serious substance abuse.** *Am J Drug Alcohol Abuse* 1997, **23:**343-54.
15. Rush B, Ellis K, Crowe T, Powell L: **How general practitioners view alcohol use. Clearing up the confusion.** *Can Fam Physician* 1994, **40:**1570-1579.
16. Makoul G, Dhurandhar A, Goel MS, Scholtens D, Rubin AS: **Communication about behavioral health risks: a study of videotaped encounters in 2 internal medicine practices.** *J Gen Intern Med* 2006, **21:**698-703.
17. McCormick KA, Cochran NE, Back AL, Merrill JO, Williams EC, Bradley KA: **How primary care providers talk to patients about alcohol: a qualitative study.** *J Gen Intern Med* 2006, **21:**966-72.

18. Adams A, Ockene JK, Wheller EV, Hurley TG: **Alcohol counseling: physicians will do it.** *J Gen Intern Med* 1998, **13:**692-98.
19. Ockene JK, Wheeler EV, Adams A, Hurley TG, Hebert J: **Provider training for patient-centered alcohol counseling in a primary care setting.** *Arch Intern Med* 1997, **157:**2334-41.
20. Saitz R, Sullivan LM, Samet JH: **Training community-based clinicians in screening and brief intervention for substance abuse problems: Translating evidence into practice.** *Subst Abus* 2000, **21:**21-31.
21. D'Onofrio G, Nadel ES, Degutis LC, Sullivan LM, Casper K, Bernstein E, Samet JH: **Improving emergency medicine residents' approach to patients with alcohol problems: a controlled educational trial.** *Ann Emerg Med* 2002, **40:**50-62.
22. Kaner EF, Lock CA, McAvoy BR, Heather N, Gilvarry E: **A RCT of three training and support strategies to encourage implementation of screening and brief alcohol intervention by general practitioners.** *British Journal of General Practice* 1999, **49:**699-703.
23. Karam-Hage M, Nerenberg L, Brower KJ: **Modifying residents' professional attitudes about substance abuse treatment and training.** *Am J Addict* 2001, **10:**40-47.
24. Rose AJ, Stein MR, Arnsten JH, Saitz R: **Teaching internal medicine resident physicians about alcoholics anonymous:a pilot study of an educational intervention.** *Subst Abus* 2006, **27:**5-11.
25. Friedmann PD, Rose J, Hayaki J, Ramsey S, Charuvastra A, Dube C, Herman D, Stein MD: **Training primary care clinicians in maintenance care for moderated alcohol use.** *J Gen Intern Med* 2006, **21:**1269-75.
26. Levin FR, Owen P, Stinchfield R, Rabinowitz E, Pace N: **Use of standardized patients to evaluate the physicians in residence program: a substance abuse training approach.** *J Addict Dis* 1999, **18:**39-50.
27. Humair JP, Cornuz J: **A new curriculum using active learning methods and standardized patients to train residents in smoking cessation.** *J Gen Intern Med* 2003, **18:**1023-27.
28. Ury WA, Rahn M, Tolentino V, Pignotti MG, Yoon J, McKegney P, Sulmasy DP: **Can a pain management and palliative care curriculum improve the opioid prescribing practices of medical residents?** *J Gen Intern Med* 2002, **17:**625-31.
29. Amodeo M, Litchfield L: **Integrating Substance Abuse Content into Social Work Courses: Effects of Intensive Faculty Training.** *Subst Abus* 1999, **20:**5-16.
30. Fleming M, Barry K, Davis A, Kropp S, Kahn R, Rivo M: **Medical education about substance abuse: changes in curriculum and faculty between 1976 and 1992.** *Acad Med* 1994, **69:**362-69.
31. Lewis DC, Niven RG, Czechowicz D, Trumble JG: **A review of medical education in alcohol and other drug abuse.** *JAMA* 1987, **257:**2945-48.
32. Alford DP, Clark TW, Samet JH: **Promoting substance abuse education among generalists: the Chief Resident Immersion Training (CRIT) program.** *Journal of General Internal Medicine* 2004, **19(Suppl):**93.
33. Gual A, Anderson P, Sequra L, Colom J: **Alcohol and Primary Health Care: Training Programme on Identification and Brief Interventions.** Barcelona, Spain, Department of Health of the Government of Catalonia; 2005.
34. Isaacson JH, Fleming M, Kraus M, Kahn R, Mundt M: **A national survey of training in substance use disorders in residency programs.** *J Stud Alcohol* 2000, **61:**912-15.
35. Spencer JA, Jordan RK: **Learner centred approaches in medical education.** *BMJ* 1999, **318:**1280-1283.
36. Brown BS: **Reducing impediments to technology transfer in drug abuse programming.** *NIDA Res Monogr* 1995, **155:**169-85.
37. Curran VR, Fleet L: **A review of evaluation outcomes of web-based continuing medical education.** *Med Educ* 2005, **39:**561-67.
38. Cook DA, Dupras DM, Thompson WG: **An online core curriculum in primary care medicine for internal medicine residents.** *Med Educ* 2003, **37:**1043.
39. Wellbery C, Gooch R: **A Web-based multimedia medical humanities curriculum.** *Fam Med* 2005, **37:**165-67.
40. Peterson MW, Galvin JR, Dayton C, D'Alessandro MP: **Realizing the promise: delivering pulmonary continuing medical education over the Internet.** *Chest* 1999, **115:**1429-36.
41. Godin P, Hubbs R, Woods B, Tsai M, Nag D, Rindfleish T, Dev P, Melmon KL: **New paradigms for medical decision support and education: the Stanford Health Information Network for Education.** *Top Health Inf Manage* 1999, **20:**1-14.
42. Bennett NL, Casebeer LL, Kristofco RE, Strasser SM: **Physicians' Internet information-seeking behaviors.** *J Contin Educ Health Prof* 2004, **24:**31-38.
43. Sargeant J, Curran V, Jarvis-Selinger S, Ferrier S, Allen M, Kirby F, Ho K: **Interactive on-line continuing medical education: physicians' perceptions and experiences.** *J Contin Educ Health Prof* 2004, **24:**227-36.
44. Fordis M, King JE, Ballantyne CM, Jones PH, Schneider KH, Spann SJ, Greenberg SB, Greisinger AJ: **Comparison of the instructional efficacy of Internet-based CME with live interactive CME workshops: a randomized controlled trial.** *JAMA* 2005, **294:**1043-51.
45. Proude EM, Conigrave KM, Haber PS: **Effectiveness of skills-based training using the Drink-less package to increase family practitioner confidence in intervening for alcohol use disorders.** *BMC Med Educ* 2006, **6:**8.
46. Casebeer L, Kristofco RE, Strasser S, Reilly M, Krishnamoorthy P, Rabin A, Zheng S, Karp S, Myers L: **Standardizing evaluation of on-line continuing medical education: physician knowledge, attitudes, and reflection on practice.** *J Contin Educ Health Prof* 2004, **24:**68-75.
47. Short LM, Surprenant ZJ, Harris JM Jr: **A community-based trial of an online intimate partner violence CME program.** *Am J Prev Med* 2006, **30:**181-85.
48. Harris JM Jr, Kutob RM, Surprenant ZJ, Maiuro RD, Delate TA: **Can Internet-based education improve physician confidence in dealing with domestic violence?** *Fam Med* 2002, **34:**287-92.
49. Wong G, Greenhalgh T, Russell J, Boynton P, Toon P: **Putting your course on the Web: lessons from a case study and systematic literature review.** *Med Educ* 2003, **37:**1020-1023.
50. National Institute on Alcohol Abuse and Alcoholism: *U.S. Department of Health and Human Services. Helping Patients Who Drink Too Much: A Clinician's Guide. Updated 2005 Edition. NIH Publication No. 07-3769* Rockville, MD; 2007.
51. **Alcohol Screening and Brief Intervention Curriculum** [http://www.mdalcoholtraining.org]
52. National Institute on Alcohol Abuse and Alcoholism: *Helping Patients with Alcohol Problems. A Health Practitioner's Guide* Bethesda, MD: National Institutes of Health; 2003.
53. **SAS/STAT Software, Version 8.2 of the SAS System for Windows.** SAS Institute Inc. Cary, NC; 1999.
54. Stimmel B, Cohen D, Colliver J, Swartz M: **An Assessment of House Staff's Knowledge of Alcohol and Substance Abuse Utilizing Standardized Patients.** *Subst Abus* 2000, **21:**1-7.
55. Slotnick HB: **How doctors learn: education and learning across the medical-school-to-practice trajectory.** *Acad Med* 2001, **76:**1013-26.
56. Rogers EM: *Diffusion of Innovations. New York, NY* 5th edition. 2003.
57. Fleming M, Murray M: *A Medical Education Model for the Prevention and Treatment of Alcohol Use Disorders* National Institute of Alcohol Abuse and Alcoholism; 1998.
58. Sirica C, Lewis DC: **Training about alcohol and substance abuse for all primary care physicians.** *A conference sponsored by the Josiah Macy, Jr. Foundation, Phoenix, Arizona, Josiah Macy, Jr. Foundation, New York, NY* 1994.
59. Murray M, Fleming M: **Prevention and treatment of alcohol-related problems: an international medical education model.** *Acad Med* 1996, **71:**1204-10.
60. Harris JM, Salasche SJ, Harris RB: **Can Internet-based continuing medical education improve physicians' skin cancer knowledge and skills?** *J Gen Intern Med* 2001, **16:**50-56.
61. Allison JJ, Kiefe CI, Wall T, Casebeer L, Ray MN, Spettell CM, Hook EW III, Oh MK, Person SD, Weissman NW: **Multicomponent Internet continuing medical education to promote chlamydia screening.** *Am J Prev Med* 2005, **28:**285-90.
62. Stewart M, Marshall JN, Ostbye T, Feightner JW, Brown JB, Harris S, Galajda J: **Effectiveness of case-based on-line learning of evidence-based practice guidelines.** *Fam Med* 2005, **37:**131-38.
63. **ProjectCork** [http://www.projectcork.org/]
64. **Alcohol Research and Education** [http://webapps.health.ufl.edu/aec/jsp/02CourseList.jsp]. Division of Addiction Medicine, University of Florida
65. Garnick DW, Lee MT, Chalk M, Gastfriend D, Horgan CM, McCorry F, McLellan AT, Merrick EL: **Establishing the feasibility of per-**

formance measures for alcohol and other drugs. *J Subst Abuse Treat* 2002, **23:**375-85.
66. Garnick DW, Horgan CM, Chalk M: **Performance measures for alcohol and other drug services.** *Alcohol Res Health* 2006, **29:**19-26.

Pre-publication history

The pre-publication history for this paper can be accessed here:

http://www.biomedcentral.com/1472-6920/8/11/prepub

Ankle Brachial Index Combined With Framingham Risk Score to Predict Cardiovascular Events and Mortality

A Meta-analysis

Ankle Brachial Index Collaboration

Context Prediction models to identify healthy individuals at high risk of cardiovascular disease have limited accuracy. A low ankle brachial index (ABI) is an indicator of atherosclerosis and has the potential to improve prediction.

Objective To determine if the ABI provides information on the risk of cardiovascular events and mortality independently of the Framingham risk score (FRS) and can improve risk prediction.

Data Sources Relevant studies were identified. A search of MEDLINE (1950 to February 2008) and EMBASE (1980 to February 2008) was conducted using common text words for the term *ankle brachial index* combined with text words and Medical Subject Headings to capture prospective cohort designs. Review of reference lists and conference proceedings, and correspondence with experts was conducted to identify additional published and unpublished studies.

Study Selection Studies were included if participants were derived from a general population, ABI was measured at baseline, and individuals were followed up to detect total and cardiovascular mortality.

Data Extraction Prespecified data on individuals in each selected study were extracted into a combined data set and an individual participant data meta-analysis was conducted on individuals who had no previous history of coronary heart disease.

Results Sixteen population cohort studies fulfilling the inclusion criteria were included. During 480 325 person-years of follow-up of 24 955 men and 23 339 women, the risk of death by ABI had a reverse J-shaped distribution with a normal (low risk) ABI of 1.11 to 1.40. The 10-year cardiovascular mortality in men with a low ABI (≤0.90) was 18.7% (95% confidence interval [CI], 13.3%-24.1%) and with normal ABI (1.11-1.40) was 4.4% (95% CI, 3.2%-5.7%) (hazard ratio [HR], 4.2; 95% CI, 3.3-5.4). Corresponding mortalities in women were 12.6% (95% CI, 6.2%-19.0%) and 4.1% (95% CI, 2.2%-6.1%) (HR, 3.5; 95% CI, 2.4-5.1). The HRs remained elevated after adjusting for FRS (2.9 [95% CI, 2.3-3.7] for men vs 3.0 [95% CI, 2.0-4.4] for women). A low ABI (≤0.90) was associated with approximately twice the 10-year total mortality, cardiovascular mortality, and major coronary event rate compared with the overall rate in each FRS category. Inclusion of the ABI in cardiovascular risk stratification using the FRS would result in reclassification of the risk category and modification of treatment recommendations in approximately 19% of men and 36% of women.

Conclusion Measurement of the ABI may improve the accuracy of cardiovascular risk prediction beyond the FRS.

JAMA. 2008;300(2):197-208 www.jama.com

MAJOR CARDIOVASCULAR and cerebrovascular events including myocardial infarction and stroke often occur in individuals without known preexisting cardiovascular disease. The prevention of such events, including the accurate identification of those at risk,[1] remains a serious public health challenge. Scoring equations to predict those at increased risk have been developed using cardiovascular risk factors, including cigarette smoking, blood pressure, total and high-density lipoprotein cholesterol, and diabetes mellitus. The Framingham risk score (FRS)[2,3] is often considered the reference standard but has limited accuracy, tending to overestimate risk in low-risk populations and underestimate risk in high-risk populations.[4] The incorporation of other risk markers, such as the metabolic syndrome[5] and plasma C-reactive protein,[6,7] has had partial success in improving prediction, and attention also is being given to indicators of asymptomatic atherosclerosis, such as coronary artery calcium, carotid intima media thickness, and the ankle brachial index (ABI).[1]

The ABI, which is the ratio of systolic pressure at the ankle to that in the arm, is quick and easy to measure and

CME available online at www.jamaarchivescme.com and questions on p 225.

A complete list of the investigators participating in the Ankle Brachial Index Collaboration appears at the end of this article.
Corresponding Author: Gerry Fowkes, PhD, Wolfson Unit for Prevention of Peripheral Vascular Diseases, Public Health Sciences, University of Edinburgh, Teviot Place, Edinburgh EH8 9AG, Scotland (Gerry.Fowkes@ed.ac.uk).

has been used for many years in vascular practice to confirm the diagnosis and assess the severity of peripheral artery disease in the legs. Most commonly the ABI is calculated by measuring the systolic blood pressure in the posterior tibial and/or the dorsalis pedis arteries either in both legs or 1 leg chosen at random (using a Doppler probe or alternative pulse sensor), with the lowest ankle pressure then divided by the brachial systolic blood pressure. In addition to peripheral artery disease, the ABI also is an indicator of generalized atherosclerosis because lower levels have been associated with higher rates of concomitant coronary and cerebrovascular disease, and with the presence of cardiovascular risk factors.[8] In population cohort studies in the United States[9-12] and Europe,[13-17] a low ABI has been related to an increased incidence of mortality (total and cardiovascular), myocardial infarction, and stroke. These increased relative risks have been shown to be independent of baseline cardiovascular disease and risk factors, suggesting that the ABI might have an independent role in predicting cardiovascular events.

The objective of our study was to determine if the ABI provides information on the risk of cardiovascular events and mortality independently of the FRS and can improve risk prediction. To enhance the representativeness of our study and to maximize participant numbers, we formed the Ankle Brachial Index Collaboration with the intent of including all major observational studies that had investigated longitudinally the ABI and incidence of cardiovascular events and mortality in general populations. At the same time we wished to identify a normal (low risk) level of the ABI that could be used in future studies and in clinical practice.

METHODS

The study design was an individual participant data meta-analysis of population-based cohort studies. The criteria for study inclusion were that the study contained participants of any age and sex derived from a general population (ie, not a specific disease group), ABI was measured at baseline using a technique standardized in each study, and individuals were followed up systematically to detect total and cardiovascular mortality.

At initial meetings of epidemiologists interested in the ABI, studies fulfilling the inclusion criteria were identified. A search was conducted of MEDLINE from 1950 to February 2008 and EMBASE from 1980 to February 2008. Reference lists and conference proceedings also were searched to identify possible additional studies. The following search terms were used: *ABPI.tw, ABI.tw, AAI.tw, ankle brachial pressure index $.tw, ankle brachial pressure$.tw, ankle brachial index$.tw. (or ankle brachial index/), ankle arm index$.tw, ankle arm blood pressure$.tw, ankle arm blood pressure index$.tw, ankle blood pressure$.tw, follow up stud$.tw, follow up studies/ or follow up/, epidemiological stud$.tw, epidemiological studies/ or epidemiology/, cohort$.tw, cohort analysis/ or cohort studies/.*

Further studies and unpublished data were sought by discussion between collaborators, cardiovascular epidemiologists, and vascular physicians and by correspondence with the Asia Pacific Cohort Studies Collaboration. Possible studies for inclusion were independently assessed for suitability by 2 collaborators (G.F. and J.P.) and any lack of clarity or disagreement was resolved by discussion.

The principal authors or lead investigators of studies were invited to join the ABI Collaboration and, following acceptance, were sent a questionnaire enquiring about the availability of specific study data. On reviewing responses to these questionnaires, a set of data that were commonly available was agreed on, and each study transferred their relevant data to the coordinating center.

Requested data included individual demographic characteristics (eg, sex, age, height, and weight), baseline clinical cofactors (eg, systolic and diastolic blood pressure, cholesterol, diabetes, and cigarette smoking), details of baseline ABI measurements, and information on nonfatal and fatal events during follow-up. For these analyses, the participants included had no previous history of coronary heart disease (CHD) as defined in each study, a value for ABI recorded at baseline, and follow-up dates or times to events. Data from collaborators were extracted and analyzed using SPSS version 14 (SPSS Inc, Chicago, Illinois) and SAS version 9.1 (SAS Institute Inc, Cary, North Carolina).

A FRS was derived for each individual using the sex-specific prediction formulas proposed by Wilson et al[3] based on conventional cardiovascular risk factors (age, total and high-density lipoprotein cholesterol categories, blood pressure categories, diabetes, and smoking status). When data on some of the variables necessary to calculate the FRS were incomplete, missing values, amounting to 3.9% of total values, were imputed using the expectation-maximization procedure for multivariate normal data, which is implemented in SPSS.

Overall (all studies combined) hazard ratios (HRs) for ABI, subdivided into 10 categories compared with a reference range of 1.11 to 1.20, were obtained for men and women for each of 3 outcomes of total mortality, cardiovascular mortality, and major coronary events (ie, coronary death, nonfatal myocardial infarction), and patterns of risk examined. Coronary revascularization and angina were not included as end points. The HRs for low vs normal ABI, which was categorized into 4 groups for the 3 outcomes of total mortality, cardiovascular mortality, and major coronary events were obtained from a proportional hazards model stratified by sex and study, both unadjusted and adjusted for FRS (categorized into 5 strata for men and 4 for women). These HRs were then pooled using a random-effects model and summarized using forest plots (Review Manager version 4.2.9, Cochrane Collaboration, Oxford, England).

Kaplan-Meier estimates and standard errors for outcome rates (total mortality, cardiovascular mortality, and major coronary events) at 10 years were obtained for each study stratified by sex and categories for FRS and ABI. Out-

come rates for studies within strata were combined to provide overall summaries using random-effects pooling.[18] Area under receiver operating characteristic curves were calculated for the prediction of events using the FRS alone and with the addition of the ABI.

RESULTS

The literature search and information from experts identified 1075 citations from which 20 studies that fulfilled the inclusion criteria were identified (FIGURE 1). Selected investigators from 16 of these studies[9-17,19-25] agreed to participate in the ABI Collaboration and provided data prior to the analysis. The participating studies and investigators are listed at the end of this article. The studies were based in Australia, Belgium, Italy, Netherlands, Sweden, the United Kingdom, and the United States and comprised predominantly white populations except for the Honolulu Heart Program (Japanese Americans)[11] and the Strong Heart Study (American Indians).[12] The populations in the Cardiovascular Health Study[10] and the Atherosclerosis Risk in Communities Study[9] comprised 15% and 26% blacks, respectively. In the San Luis Valley Diabetes Study,[24] the included healthy population without diabetes was 42% Hispanic. Eleven studies included both sexes, 4 included only men, and 1 included only women.

The characteristics of the participants in the studies at baseline when the ABI was measured are shown in TABLE 1. A total of 24 955 men and 23 339 women without a history of CHD were included. They were late middle aged to elderly with a mean age in the studies ranging from 47 to 78 years. The 10-year mean (SD) incidence of CHD predicted by the FRS at baseline varied across studies from 11.0% (6.1%) to 31.6% (14.1%) in men and from 7.1% (6.1%) to 14.5% (10.1%) in women. The mean (SD) ABI was greater than 1.00 in all studies and ranged from 1.02 (0.13) to 1.21 (0.13) in men and 1.01 (0.16) to 1.15 (0.17) in women; most of the studies comprising both sexes had higher mean values in men than in women, as previously reported.[24]

TABLE 2 and TABLE 3 show the total mortality, cardiovascular mortality, and major coronary events occurring during follow-up in each of the studies for men and women, respectively. Median duration of follow-up ranged from 3 to

Figure 1. Flow Diagram of Selection of Studies for Inclusion in Meta-analysis

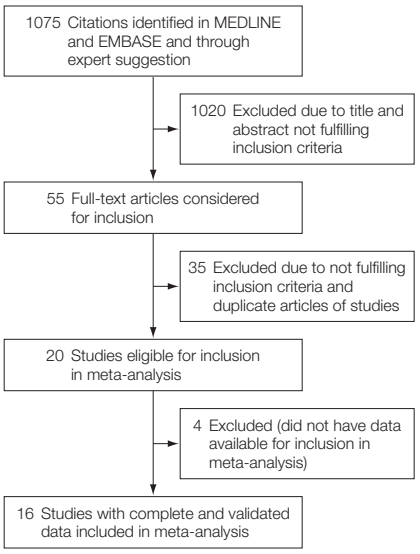


Table 1. Baseline Characteristics of Individuals in Studies in the Ankle Brachial Index (ABI) Collaboration

Source	Study	No. of Individuals[a]		Mean (SD)				
					FRS, %[b]		ABI	
		Men (n = 24 955)	Women (n = 23 339)	Age, y	Men	Women	Men	Women
Weatherley et al,[9] 2007	ARIC	6105	8004	54 (5.7)	12.8 (7.6)	7.3 (6.0)	1.17 (0.13)	1.12 (0.13)
Kornitzer et al,[17] 1995	Belgian Physical Fitness	2068	0	47 (4.4)	11.0 (6.1)	NA	1.21 (0.13)	NA
Newman et al,[10] 1999	Cardiovascular Health	1846	2779	73 (5.5)	25.4 (12.5)	8.0 (5.3)	1.10 (0.19)	1.06 (0.15)
Leng et al,[13] 1996	Edinburgh Artery	690	702	64 (5.7)	26.2 (13.0)	11.5 (6.2)	1.07 (0.19)	1.01 (0.16)
Murabito et al,[19] 2002	Framingham Offspring	1423	1703	58 (9.6)	15.3 (10.3)	7.5 (5.9)	1.16 (0.12)	1.10 (0.10)
Fowler et al,[20] 2002	Health in Men	2771	0	72 (4.4)	29.4 (9.6)	NA	1.07 (0.17)	NA
Abbott et al,[11] 2000	Honolulu Heart Program	3123	0	78 (4.6)	31.6 (14.1)	NA	1.05 (0.17)	NA
Jager et al,[21] 1999	Hoorn	270	284	63 (7.2)	26.8 (13.9)	14.5 (10.1)	1.03 (0.14)	1.02 (0.12)
McDermott et al,[22] 2004	InCHIANTI	481	569	67 (15.5)	24.8 (15.4)	8.0 (5.8)	1.04 (0.16)	1.05 (0.14)
Hooi et al,[14] 2004	Limburg PAOD	1031	1320	57 (9.4)	20.2 (10.6)	11.7 (5.8)	1.08 (0.16)	1.07 (0.13)
Ogren et al,[15] 1993	Men Born in 1914	391	0	69 (0.5)	31.5 (10.5)	NA	1.02 (0.13)	NA
Van der Meer et al,[16] 2004	Rotterdam	2134	3515	69 (9.2)	29.6 (15.6)	10.2 (7.2)	1.10 (0.21)	1.05 (0.21)
Criqui et al,[23] 1992	San Diego	244	314	66 (10.4)	21.6 (12.9)	7.8 (5.1)	1.08 (0.19)	1.02 (0.12)
Hiatt et al,[24] 1995	San Luis Valley Diabetes	674	838	53 (12.1)	15.6 (12.0)	9.1 (9.4)	1.16 (0.15)	1.10 (0.14)
Resnick et al,[12] 2004	Strong Heart	1704	2622	56 (8.0)	15.5 (9.6)	10.8 (7.3)	1.15 (0.14)	1.15 (0.17)
McDermott et al,[25] 2000	Women's Health and Aging	0	689	78 (8.1)	NA	7.1 (6.1)	NA	1.05 (0.21)

Abbreviations: ARIC, Atherosclerosis Risk in Communities; FRS, Framingham risk score; InCHIANTI, Invecchiare in Chianti; NA, not applicable; PAOD, peripheral arterial occlusive disease.
[a] No history of coronary heart disease (including myocardial infarction, angina, and revascularization as defined in each study), ABI available at baseline, and follow-up data available.
[b] Predicted percentage at 10 years for incidence of coronary heart disease, including coronary death, myocardial infarction, and angina.

16.7 years, with 9 of the 16 studies having more than 10 years of follow-up. Overall, 9924 deaths occurred during 480 325 person-years of follow-up with around one-quarter of deaths due to CHD or stroke in both men and women. The annual rates of deaths and events varied considerably between the studies. For example, men in the Belgian Physical Fitness Study had a mean (SD) age of 47 (4.4) years and the annual mortality was 0.37% (95% confidence interval [CI], 0.29%-0.45%), whereas men in the Honolulu Heart Program had a mean (SD) age of 78 (4.6) years and the annual mortality was 4.91% (95% CI, 4.59%-5.22%) (Table 2). Likewise, for women annual mortality varied between 0.55% (95% CI, 0.42%-0.68%) in the Framingham Offspring Study and 7.34% (95% CI, 6.39%-8.29%) in the Women's Health and Aging Study (Table 3).

The HRs for death for different levels of ABI compared with a reference ABI of 1.11 to 1.20 in all studies combined formed a reverse J-shaped curve for both men and women (FIGURE 2). For levels of ABI below 1.11, the HRs increased consistently with decreasing ABI. For an ABI of greater than 1.40, the HRs also were increased in men (1.38; 95% CI, 1.17-1.62) and in women (1.23; 95% CI, 1.00-1.52). For levels of ABI from 1.11 to 1.40, only small and mostly nonsignificant differences in HRs were found. TABLE 4 and TABLE 5 show the HRs for total and cardiovascular mortality and major coronary events by ABI in men and women, respectively. The patterns of risk for cardiovascular mortality and major coronary events were similar to that for total mortality; for levels of ABI below 1.11, the HRs for cardiovascular mortality were consistently higher than for total mortality.

Values of the ABI less than 0.90 have been taken traditionally as a measure of increased risk. In nearly all the studies in men (FIGURE 3), the HRs for total mortality were statistically significantly higher in individuals with an ABI of 0.90 or less compared with individuals with normal ABI values of 1.11 to 1.40 (HR, 3.33; 95% CI, 2.74-4.06). In women, the results were more heterogeneous (FIGURE 4), but the HR of 2.71 (95% CI, 2.03-3.62) was comparable

Table 2. Total Mortality, Cardiovascular Mortality, and Major Coronary Events for Men in Studies in the Ankle Brachial Index Collaboration

		Total Mortality			Cardiovascular Mortality[a]			Major Coronary Events[b]		
Study	Follow-up, Median (IQR), y	Person-Years of Follow-up (n = 233 457)	No. of Deaths (n = 5582)	Annual Mortality, % (95% CI)	Person-Years of Follow-up (n = 233 457)	No. of Deaths (n = 1507)	Annual Mortality, % (95% CI)	Person-Years of Follow-up (n = 205 628)	No. of Events (n = 2255)	Annual Events, % (95% CI)
ARIC[9]	13.1 (12.4-13.9)	76 497	903	1.18 (1.10-1.26)	76 497	170	0.22 (0.19-0.26)	73 991	571	0.77 (0.71-0.83)
Belgian Physical Fitness[17]	10.9 (10.5-11.4)	22 292	83	0.37 (0.29-0.45)	22 292	13	0.06 (0.03-0.09)	22 136	98	0.44 (0.36-0.53)
Cardiovascular Health[10]	11.0 (7.2-11.6)	16 583	839	5.06 (4.73-5.39)	16 583	263	1.59 (1.40-1.78)	15 542	432	2.78 (2.52-3.04)
Edinburgh Artery[13]	15.5 (9.0-15.9)	8667	295	3.40 (3.02-3.79)	8667	84	0.97 (0.76-1.18)	8090	113	1.40 (1.14-1.65)
Framingham Offspring[19]	7.4 (6.6-8.2)	10 182	113	1.11 (0.91-1.31)	10 182	20	0.20 (0.11-0.28)	10 052	56	0.56 (0.41-0.70)
Health in Men[20]	6.3 (5.9-6.5)	16 446	402	2.44 (2.21-2.68)	16 446	114	0.69 (0.57-0.82)	NA	NA	NA
Honolulu Heart Program[11]	6.2 (5.5-6.9)	17 976	882	4.91 (4.59-5.22)	17 976	231	1.29 (1.12-1.45)	17 703	205	1.16 (1.00-1.32)
Hoorn[21]	12.5 (9.8-13.1)	2969	88	2.96 (2.35-3.57)	2969	26	0.88 (0.54-1.21)	NA	NA	NA
InCHIANTI[22]	3.0 (2.9-3.1)	1427	30	2.10 (1.36-2.85)	1427	11	0.77 (0.32-1.22)	NA	NA	NA
Limburg PAOD[14]	7.1 (6.6-7.7)	7088	148	2.09 (1.76-2.42)	7088	34	0.48 (0.32-0.64)	6864	82	1.19 (0.94-1.45)
Men Born in 1914[15]	13.3 (8.1-13.9)	4248	182	4.28 (3.68-4.89)	4248	70	1.65 (1.26-2.03)	4028	92	2.28 (1.82-2.75)
Rotterdam[16]	10.9 (8.2-11.8)	20 538	813	3.96 (3.70-4.23)	20 538	221	1.08 (0.94-1.23)	19 805	260	1.31 (1.15-1.47)
San Diego[23]	16.7 (10.4-22.3)	3843	156	4.06 (3.44-4.68)	3843	77	2.00 (1.56-2.45)	3581	80	2.23 (1.75-2.72)
San Luis Valley Diabetes[24]	15.6 (14.4-16.9)	9765	167	1.71 (1.45-1.97)	9765	51	0.52 (0.38-0.67)	9265	82	0.89 (0.69-1.08)
Strong Heart[12]	9.7 (8.9-10.4)	14 935	481	3.22 (2.94-3.50)	14 935	122	0.82 (0.67-0.96)	14 573	184	1.27 (1.09-1.46)

Abbreviations: ARIC, Atherosclerosis Risk in Communities; CI, confidence interval; InCHIANTI, Invecchiare in Chianti; IQR, interquartile range; NA, data not available; PAOD, peripheral arterial occlusive disease.
[a] Defined as death due to coronary heart disease or stroke.
[b] Defined as myocardial infarction or deaths from coronary heart disease.

Table 3. Total Mortality, Cardiovascular Mortality, and Major Coronary Events for Women in Studies in the Ankle Brachial Index Collaboration

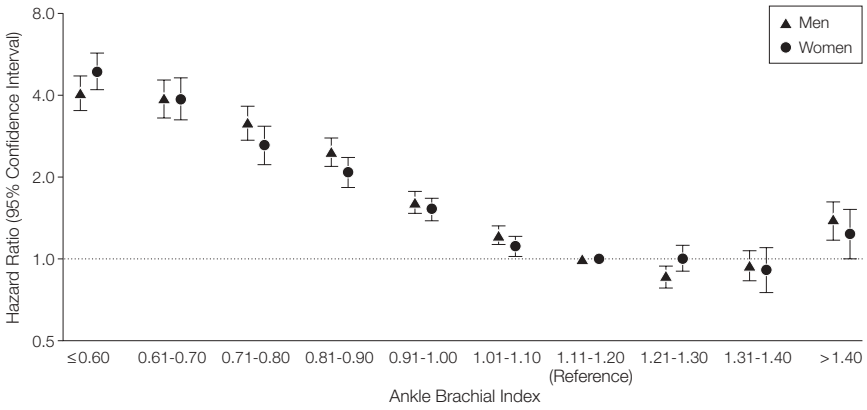
		Total Mortality			Cardiovascular Mortality[a]			Major Coronary Events[b]		
Study	Follow-up, Median (IQR), y	Person-Years of Follow-up (n = 246 868)	No. of Deaths (n = 4342)	Annual Mortality, % (95% CI)	Person-Years of Follow-up (n = 246 868)	No. of Deaths (n = 1211)	Annual Mortality, % (95% CI)	Person-Years of Follow-up (n = 238 066)	No. of Events (n = 1629)	Annual Events, % (95% CI)
ARIC[9]	13.2 (12.4-13.9)	102 458	773	0.75 (0.70-0.81)	102 458	133	0.13 (0.11-0.15)	101 121	362	0.36 (0.32-0.39)
Cardiovascular Health[10]	11.2 (8.3-11.6)	27 447	851	3.10 (2.90-3.31)	27 447	262	0.95 (0.84-1.07)	26 652	374	1.40 (1.26-1.54)
Edinburgh Artery[13]	15.8 (14.2-16.1)	9836	200	2.03 (1.75-2.31)	9836	41	0.42 (0.29-0.54)	9602	57	0.59 (0.44-0.75)
Framingham Offspring[19]	7.4 (6.6-8.3)	12 344	68	0.55 (0.42-0.68)	12 344	5	0.04 (0.01-0.08)	12 272	24	0.20 (0.12-0.27)
Hoorn[21]	12.6 (10.6-13.2)	3212	76	2.37 (1.84-2.89)	3212	23	0.72 (0.42-1.01)	NA	NA	NA
InCHIANTI[22]	3.0 (2.9-3.2)	1711	26	1.52 (0.94-2.10)	1711	12	0.70 (0.31-1.10)	NA	NA	NA
Limburg PAOD[14]	7.1 (6.7-7.6)	9273	114	1.23 (1.01-1.45)	9273	26	0.28 (0.17-0.39)	9168	53	0.58 (0.42-0.73)
Rotterdam[16]	11.1 (9.3-12.1)	35 407	1131	3.19 (3.01-3.38)	35 407	352	0.99 (0.89-1.10)	34 968	283	0.81 (0.72-0.90)
San Diego[23]	19.6 (13.0-22.6)	5443	177	3.25 (2.78-3.72)	5443	76	1.40 (1.08-1.71)	5361	65	1.21 (0.92-1.51)
San Luis Valley Diabetes[24]	15.8 (14.6-17.1)	12 542	163	1.30 (1.10-1.50)	12 542	53	0.42 (0.31-0.54)	12 293	58	0.47 (0.35-0.59)
Strong Heart[12]	9.9 (9.1-10.7)	24 305	551	2.27 (2.08-2.45)	24 305	137	0.56 (0.47-0.66)	24 010	183	0.76 (0.65-0.87)
Women's Health and Aging[25]	5.0 (3.8-5.1)	2890	212	7.34 (6.39-8.29)	2890	91	3.15 (2.51-3.79)	2620	170	6.49 (5.55-7.43)

Abbreviations: ARIC, Atherosclerosis Risk in Communities; CI, confidence interval; InCHIANTI, Invecchiare in Chianti; IQR, interquartile range; NA, data not available; PAOD, peripheral arterial occlusive disease.
[a] Defined as death due to coronary heart disease or stroke.
[b] Defined as myocardial infarction or deaths from coronary heart disease.

with that in men. Likewise, significantly increased HRs were found in men and in women both for cardiovascular mortality (men: 4.21 [95% CI, 3.29-5.39]; women: 3.46 [95% CI, 2.36-5.08]), and for major coronary events (men: 2.97 [95% CI, 2.33-3.78]; women: 3.05 [95% CI, 2.25-4.15]). Adjustment of the HRs for individuals with an ABI of 0.90 or less relative to an ABI of 1.11 to 1.40 for FRS reduced the HRs but they were still elevated substantially and significantly. The adjusted HRs for total mortality were 2.34 (95% CI, 1.97-2.78) in men vs 2.35 (95% CI, 1.76-3.13) in women; cardiovascular mortality, 2.92 (95% CI, 2.31-3.70) in men vs 2.97 (95% CI, 2.02-4.35) in women; and major coronary events, 2.16 (95% CI, 1.76-2.66) in men vs 2.49 (95% CI, 1.84-3.36) in women.

TABLE 6 and TABLE 7 show the effect of inclusion of an ABI measurement on the apparent risk of 10-year total mortality, cardiovascular mortality, and major coronary events over the range of FRS categories in men and women, respectively. Compared with the overall rates without ABI included, an ABI of 0.90 or less was associated with a greatly increased risk of mortality (total and cardiovascular) and major coronary events across all FRS categories in both men and women, but more so in the

Figure 2. Hazard Ratios for Total Mortality in Men and Women by Ankle Brachial Index at Baseline for All Studies Combined in the ABI Collaboration



Hazard ratios are not adjusted for age or cardiovascular risk factors.

lower than in the higher FRS categories. Women had especially high mortality and event rates in the lowest FRS category. Men and women with an ABI from 0.91 to 1.10 also had higher mortality and event rates compared with those with a normal ABI (1.11-1.40) but the magnitudes of the increase were much less than for those with an ABI of 0.90 or less. Those with an ABI greater than 1.40 also had higher rates across most FRS categories.

Inclusion of the ABI had an overall effect on the prediction of events, especially in women. When predicting major coronary events using only the FRS, the area under the receiver operating characteristic curve was 0.646 (95% CI, 0.643-0.657) and with the addition of the ABI was 0.655 (95% CI, 0.643-0.666) in men vs 0.605 (95% CI,

Table 4. Hazard Ratios (HRs) for Total Mortality, Cardiovascular Mortality, and Major Coronary Events by Ankle Brachial Index (ABI) at Baseline for Men in All Studies Combined in the ABI Collaboration

	ABI									
	≤0.60	0.61-0.70	0.71-0.80	0.81-0.90	0.91-1.00	1.01-1.10	1.11-1.20	1.21-1.30	1.31-1.40	>1.40
Sample size (n = 24 955)	360	279	428	774	2438	5775	7576	4936	1681	708
Total Mortality										
No. of deaths (n = 5582)	215	170	217	355	741	1338	1364	745	270	167
HR (95% CI)[a]	4.06 (3.51-4.70)	3.88 (3.30-4.55)	3.15 (2.73-3.64)	2.47 (2.19-2.78)	1.61 (1.47-1.77)	1.22 (1.13-1.32)	1 [Reference]	0.86 (0.78-0.94)	0.94 (0.83-1.07)	1.38 (1.17-1.62)
Cardiovascular Mortality[b]										
No. of deaths (n = 1507)	80	54	81	116	208	352	341	179	62	34
HR (95% CI)[a]	5.58 (4.36-7.15)	4.60 (3.44-6.14)	4.49 (3.51-5.74)	3.03 (2.45-3.75)	1.68 (1.40-2.00)	1.24 (1.07-1.44)	1 [Reference]	0.85 (0.71-1.02)	0.93 (0.71-1.22)	1.14 (0.80-1.63)
Major Coronary Events (n = 21 433)[c]										
No. of events (n = 2255)	70	48	74	119	252	516	642	353	125	56
HR (95% CI)[a]	3.45 (2.68-4.43)	2.71 (2.01-3.64)	2.76 (2.16-3.52)	2.15 (1.76-2.63)	1.43 (1.23-1.66)	1.12 (1.00-1.26)	1 [Reference]	0.78 (0.68-0.88)	0.78 (0.64-0.95)	0.90 (0.68-1.18)

Abbreviation: CI, confidence interval.
[a]The HRs are not adjusted for age or cardiovascular risk factors.
[b]Defined as death due to coronary heart disease or stroke.
[c]Defined as myocardial infarction or deaths from coronary heart disease.

Table 5. Hazard Ratios (HRs) for Total Mortality, Cardiovascular Mortality, and Major Coronary Events by Ankle Brachial Index (ABI) at Baseline for Women in All Studies Combined in the ABI Collaboration

	ABI									
	≤0.60	0.61-0.70	0.71-0.80	0.81-0.90	0.91-1.00	1.01-1.10	1.11-1.20	1.21-1.30	1.31-1.40	>1.40
Sample size (n = 23 339)	314	251	403	933	3186	6586	6862	3363	932	509
Total Mortality										
No. of deaths (n = 4342)	199	145	174	326	707	1078	999	489	125	100
HR (95% CI)[a]	4.89 (4.19-5.71)	3.88 (3.25-4.63)	2.61 (2.22-3.08)	2.08 (1.83-2.36)	1.52 (1.38-1.67)	1.11 (1.02-1.21)	1 [Reference]	1.00 (0.90-1.12)	0.91 (0.75-1.10)	1.23 (1.00-1.52)
Cardiovascular Mortality[b]										
No. of deaths (n = 1211)	79	51	66	114	218	271	241	119	24	28
HR (95% CI)[a]	7.04 (5.43-9.12)	5.06 (3.72-6.87)	3.65 (2.77-4.81)	2.77 (2.21-3.47)	1.84 (1.53-2.22)	1.14 (0.95-1.36)	1 [Reference]	1.04 (0.83-1.29)	0.74 (0.49-1.13)	1.48 (1.00-2.21)
Major Coronary Events (n = 22 486)[c]										
No. of events (n = 1629)	79	54	64	119	260	412	387	174	47	33
HR (95% CI)[a]	5.43 (4.24-6.94)	3.82 (2.86-5.11)	2.58 (1.97-3.37)	2.06 (1.67-2.53)	1.53 (1.30-1.79)	1.11 (0.97-1.28)	1 [Reference]	0.91 (0.76-1.09)	0.86 (0.64-1.17)	1.11 (0.77-1.58)

Abbreviation: CI, confidence interval.
[a]The HRs are not adjusted for age or cardiovascular risk factors.
[b]Defined as death due to coronary heart disease or stroke.
[c]Defined as myocardial infarction or deaths from coronary heart disease.

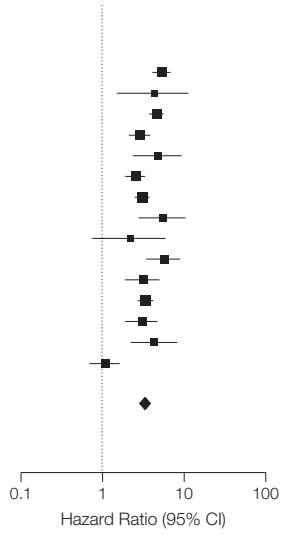
0.590-0.619) and 0.658 (95% CI, 0.644-0.672), respectively, in women.

The FRS is mostly used to predict risk of total CHD (including coronary death, myocardial infarction, and angina) and TABLE 8 shows the effect of including the ABI on this prediction. The calibration of the FRS categories was reasonable because the overall CHD rate in each FRS category was within the range predicted, except for low-risk women in which the overall CHD rate of 11% was higher than predicted. Likewise, the ability of the FRS to discriminate between risk categories was good, except that the overall CHD rate in women in the low-risk group was only slightly lower than those in the intermediate-risk group (11% vs 13%). In each category of FRS in both men and women, a low ABI (≤0.90) was associated with an increased risk of future CHD. Normal levels of the ABI (1.11-1.40) were associated with a slightly reduced risk from the overall rates but levels greater than 1.40 did not differ consistently from the overall rates, al-

Figure 3. Random Hazard Ratios for Total Mortality for Low (≤0.90) Compared With Normal (1.11-1.40) Ankle Brachial Index (ABI) in Men in Studies in the ABI Collaboration

Study	No. of Men		No. of Deaths		Hazard Ratio (95% CI)
	ABI ≤0.90	ABI, 1.11-1.40	ABI ≤0.90	ABI, 1.11-1.40	
ARIC,[9] 2007	129	4446	64	568	5.23 (4.04-6.77)
Belgian Physical Fitness,[17] 1995	25	1544	4	64	4.15 (1.52-11.37)
Cardiovascular Health,[10] 1999	220	1005	177	330	4.48 (3.72-5.39)
Edinburgh Artery,[13] 1996	94	316	69	111	2.84 (2.10-3.84)
Framingham Offspring,[19] 2002	38	1120	9	72	4.67 (2.33-9.34)
Health in Men,[20] 2002	288	1221	76	143	2.50 (1.89-3.30)
Honolulu Heart Program,[11] 2000	391	1048	204	226	3.00 (2.48-3.63)
Hoorn,[21] 1999	30	83	20	16	5.34 (2.74-10.39)
InCHIANTI,[22] 2004	53	151	7	8	2.11 (0.75-5.93)
Limburg PAOD,[14] 2004	89	459	31	36	5.50 (3.40-8.90)
Men Born in 1914,[15] 1993	46	79	35	32	3.06 (1.89-4.96)
Rotterdam,[16] 2004	318	1133	200	317	3.25 (2.72-3.88)
San Diego,[23] 1992	30	103	28	62	2.97 (1.88-4.69)
San Luis Valley Diabetes,[24] 1995	15	479	10	106	4.25 (2.22-8.13)
Strong Heart,[12] 2004	75	1006	23	288	1.06 (0.69-1.62)
Overall	1841	14 193	957	2379	3.33 (2.74-4.06)

Test for heterogeneity: $\chi^2_{14} = 64.32$; $P<.001$; $I^2 = 78.2\%$
Test for overall effect: $z = 11.98$; $P<.001$

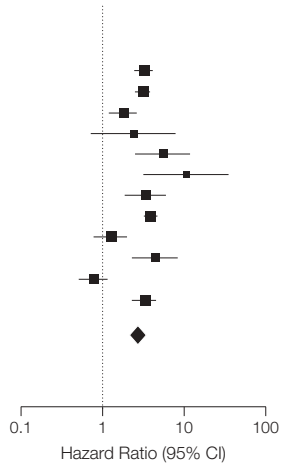


Hazard ratios are not adjusted for age or cardiovascular risk factors. Area of each square is proportional to weight of the study in the meta-analysis. ARIC indicates Atherosclerosis Risk in Communities; CI, confidence interval; InCHIANTI, Invecchiare in Chianti; PAOD, peripheral arterial occlusive disease.

Figure 4. Random Hazard Ratios for Total Mortality for Low (≤0.90) Compared With Normal (1.11-1.40) Ankle Brachial Index (ABI) in Women in Studies in the ABI Collaboration

Study	No. of Women		No. of Deaths		Hazard Ratio (95% CI)
	ABI ≤0.90	ABI, 1.11-1.40	ABI ≤0.90	ABI, 1.11-1.40	
ARIC,[9] 2007	268	4441	66	393	3.17 (2.44-4.11)
Cardiovascular Health,[10] 1999	297	1134	159	280	3.05 (2.51-3.71)
Edinburgh Artery,[13] 1996	133	192	50	46	1.77 (1.19-2.64)
Framingham Offspring,[19] 2002	46	887	3	27	2.38 (0.72-7.86)
Hoorn,[21] 1999	23	62	15	11	5.42 (2.48-11.84)
InCHIANTI,[22] 2004	37	176	8	4	10.50 (3.16-34.88)
Limburg PAOD,[14] 2004	84	486	18	33	3.34 (1.88-5.94)
Rotterdam,[16] 2004	657	1498	383	324	3.75 (3.23-4.36)
San Diego,[23] 1992	47	81	29	45	1.25 (0.78-2.00)
San Luis Valley Diabetes,[24] 1995	19	364	11	63	4.37 (2.30-8.30)
Strong Heart,[12] 2004	143	1540	24	325	0.77 (0.51-1.16)
Women's Health and Aging,[25] 2000	147	296	78	62	3.22 (2.30-4.50)
Overall	1901	11 157	844	1613	2.71 (2.03-3.62)

Test for heterogeneity: $\chi^2_{11} = 78.97$; $P<.001$; $I^2 = 86.1\%$
Test for overall effect: $z = 6.73$; $P<.001$



Hazard ratios are not adjusted for age or cardiovascular risk factors. Area of each square is proportional to weight of the study in the meta-analysis. ARIC indicates Atherosclerosis Risk in Communities; CI, confidence interval; InCHIANTI, Invecchiare in Chianti; PAOD, peripheral arterial occlusive disease.

though this may have been influenced by the relatively low numbers of participants.

The results in Table 8 also indicate in which categories of FRS the ABI is likely to change individuals' clinical risk levels (ie, between <10%, 10%-19%, and ≥20%). In men, the greatest effect would be in high-risk individuals (≥20%) with a normal ABI (1.11-1.40) in whom the risk level would be reduced to intermediate (10%-19%). All men with a low ABI (≤0.90) had a relatively high risk but their clinical risk level would not change from that predicted overall by the FRS. In women, the main effect of the ABI would be to change all women in the low FRS category (<10%) with an abnormal ABI (≤0.90 or 0.91 to 1.10 or >1.40) to a higher risk level. Also women in the intermediate FRS category (10%-19%) with a low ABI (≤0.90) would become high risk (≥20%). Table 8 also shows that the number of men changing risk category (shaded numbers) would be 4106 of 21 433 (19%) and in women would be 8154 of 22 486 women (36%).

COMMENT

Predicting future CHD and mortality accurately in individuals in the community who have no prior history of cardiovascular disease has proven difficult when based solely on traditional risk factors and scoring systems. In a recent systematic review of 27 studies using the Framingham risk equation, the predicted-to-observed ratios ranged from an underprediction of 0.43 in a high-risk population to an overprediction of 2.87 in a low-risk population.[4] We found that the ABI provided independent risk information compared with the FRS and, when combined with the FRS, a low ABI (≤0.90) approximately doubled the risk of total mortality, cardiovascular mortality, and major coronary events across all Framingham risk categories.

In predicting the 10-year risk of total CHD, our results (Table 8) indicate that approximately 1 of 5 men would have their broad category of risk (<10, 10-19, ≥20%) changed from that predicted by FRS alone to that found on inclusion of the ABI. These changes from higher to lower categories of risk would likely have an effect on decisions to commence preventive treatment, such as lipid-lowering therapy, as recommended in the Adult Treatment Panel III guidelines.[27] In contrast, the main effect in women of inclusion of the ABI would be that many at low risk with the FRS (<10%) would change to a higher risk level. In total, around 1 in 3 women would be affected. It should be recognized, however, that the proportion of men and women affected by inclusion of the ABI

Table 6. 10-Year Total Mortality, Cardiovascular Mortality, and Major Coronary Event Rates in Men by Framingham Risk Category and Ankle Brachial Index (ABI) at Baseline for All Studies Combined in the ABI Collaboration[a]

Framingham Risk Category[b]	ABI				Overall
	≤0.90	0.91-1.10	1.11-1.40	>1.40	
			Total Mortality, % (95% CI)		
1 (Lowest; n = 5746)	27.1 (16.0-38.2)	11.4 (5.9-16.8)	8.3 (5.4-11.2)	14.1 (4.2-24.0)	10.4 (6.9-13.9)
2 (n = 4319)	37.3 (17.8-56.9)	15.8 (10.6-21.0)	11.3 (8.2-14.5)	19.9 (7.5-32.4)	13.8 (9.9-17.7)
3 (n = 3544)	37.6 (26.1-49.1)	19.7 (13.6-25.9)	14.2 (9.9-18.5)	23.5 (9.5-37.6)	17.6 (13.1-22.2)
4 (n = 5814)	38.1 (28.5-47.8)	23.6 (16.9-30.4)	19.2 (14.8-23.5)	38.4 (12.3-64.6)	23.1 (17.6-28.6)
5 (Highest; n = 5532)	57.1 (45.4-68.9)	36.4 (29.1-43.7)	31.0 (25.2-36.7)	43.6 (28.1-59.1)	38.0 (30.9-45.0)
Overall (n = 24 955)	46.3 (36.1-56.6)	23.0 (15.8-30.2)	16.7 (12.4-21.0)	29.2 (18.9-39.5)	
			Cardiovascular Mortality, % (95% CI)		
1 (Lowest; n = 5746)	4.6 (0.0-10.8)	3.1 (0.0-6.5)	1.3 (0.5-2.0)	2.7 (0.0-6.8)	1.6 (0.8-2.4)
2 (n = 4319)	17.5 (6.6-28.3)	3.5 (1.5-5.5)	1.5 (0.7-2.3)	8.2 (0.0-18.8)	2.3 (1.3-3.4)
3 (n = 3544)	11.5 (2.4-20.6)	5.1 (3.1-7.2)	3.6 (1.9-5.2)	8.3 (0.3-16.2)	4.4 (2.8-6.0)
4 (n = 5814)	14.2 (10.2-18.2)	8.0 (5.2-10.8)	4.8 (3.3-6.4)	5.6 (0.0-12.2)	7.3 (5.2-9.3)
5 (Highest; n = 5532)	27.9 (20.7-35.1)	12.5 (8.9-16.1)	9.9 (6.8-13.1)	10.7 (2.0-19.4)	14.0 (10.6-17.4)
Overall (n = 24 955)	18.7 (13.3-24.1)	7.3 (5.0-9.6)	4.4 (3.2-5.7)	6.9 (2.8-11.0)	
			Major Coronary Events, % (95% CI)[c]		
1 (Lowest; n = 5643)	5.8 (0.0-12.7)	3.7 (1.4-6.0)	3.4 (2.5-4.3)	4.0 (1.1-6.8)	3.5 (2.4-4.6)
2 (n = 4151)	20.0 (9.6-30.4)	5.9 (3.6-8.1)	6.8 (5.7-8.0)	5.0 (0.7-9.3)	7.1 (5.5-8.8)
3 (n = 3241)	20.2 (8.0-32.3)	10.0 (6.2-13.8)	8.7 (6.4-11.0)	12.9 (0.0-27.8)	10.1 (7.5-12.6)
4 (n = 4179)	27.5 (18.5-36.6)	14.8 (9.9-19.7)	12.6 (9.6-15.7)	9.7 (0.0-19.7)	15.3 (11.5-19.1)
5 (Highest; n = 4219)	31.4 (21.9-40.8)	20.0 (14.4-25.5)	17.6 (12.2-23.0)	12.0 (3.6-20.3)	21.5 (16.7-26.3)
Overall (n = 21 433)	26.8 (19.5-34.1)	12.9 (9.2-16.7)	9.4 (7.4-11.4)	7.2 (4.3-10.1)	

Abbreviation: CI, confidence interval.
[a] Analysis based on random-effects pooling of Kaplan-Meier estimates from the individual studies.
[b] Categories of predicted 10-year percentage incidence of coronary heart disease, including coronary death, myocardial infarction, and angina are based on whole number cut points for scores (category 1, <10%; 2, 10%-14%; 3, 15%-19%; 4, 20%-29%; 5, ≥30%).
[c] Excludes Health in Men,[20] Hoorn,[21] and InCHIANTI[22] studies.

is approximate due to the method of estimating total CHD end points and possible residual confounding within the FRS categories.

Our results also confirm the recent findings of the Strong Heart Study,[12] Cardiovascular Health Study,[28] and Multi-ethnic Study of Atherosclerosis[29] that the relationship between ABI and cardiovascular disease is nonlinear and varies across the range of ABI. High values (>1.40) could be related to poor arterial compressibility resulting from stiffness and calcification, which may occur more commonly in those with diabetes,[29,30] and may be 1 explanation why those with an ABI greater than 1.40 are at increased risk. The differences in risk between ABI values from 1.11 to 1.40 in both men and women were so small that, for practi-

Table 7. 10-Year Total Mortality, Cardiovascular Mortality, and Major Coronary Event Rates in Women by Framingham Risk Category and Ankle Brachial Index (ABI) at Baseline for All Studies Combined in the ABI Collaboration[a]

Framingham Risk Category[b]	ABI				Overall
	≤0.90	0.91-1.10	1.11-1.40	>1.40	
			Total Mortality, % (95% CI)		
1 (Lowest; n = 5507)	44.2 (7.5-80.9)	21.3 (12.5-30.1)	14.1 (9.1-19.1)	27.4 (14.6-40.2)	18.2 (10.6-25.8)
2 (n = 6016)	28.2 (9.2-47.2)	13.3 (7.7-18.9)	10.3 (6.3-14.3)	8.1 (1.9-14.3)	12.2 (7.0-17.4)
3 (n = 5581)	27.1 (16.0-38.1)	15.2 (11.0-19.4)	10.9 (7.5-14.2)	20.6 (11.7-29.5)	15.7 (11.2-20.2)
4 (Highest; n = 6235)	31.4 (23.2-39.7)	17.6 (13.3-21.9)	16.2 (12.2-20.3)	20.9 (0.0-48.2)	19.8 (16.6-23.0)
Overall (n = 23 339)	30.1 (18.0-42.1)	16.6 (10.9-22.3)	13.1 (8.5-17.6)	26.6 (9.7-43.4)	
			Cardiovascular Mortality, % (95% CI)		
1 (Lowest; n = 5507)	45.5 (29.7-61.4)	4.5 (1.9-7.0)	4.0 (1.6-6.4)	14.1 (0.0-32.3)	4.8 (3.2-6.4)
2 (n = 6016)	15.1 (1.5-28.7)	4.1 (1.6-6.6)	2.9 (0.9-4.9)	4.3 (0.0-12.7)	3.5 (1.6-5.4)
3 (n = 5581)	9.7 (5.1-14.3)	4.4 (2.5-6.3)	3.2 (1.5-4.8)	14.7 (0.0-45.6)	4.8 (3.0-6.6)
4 (Highest; n = 6235)	15.7 (9.5-22.0)	5.1 (3.4-6.9)	5.5 (3.5-7.6)	15.5 (8.4-22.5)	6.8 (4.5-9.2)
Overall (n = 23 339)	12.6 (6.2-19.0)	4.7 (3.0-6.3)	4.1 (2.2-6.1)	6.9 (4.0-9.7)	
			Major Coronary Events, % (95% CI)[c]		
1 (Lowest; n = 5355)	29.9 (9.0-50.8)	3.9 (1.7-6.1)	5.3 (2.4-8.2)	10.7 (0.0-24.3)	5.8 (3.9-7.7)
2 (n = 5842)	16.9 (6.8-27.1)	5.1 (2.4-7.7)	3.7 (2.0-5.5)	2.1 (0.0-6.3)	4.7 (2.6-6.7)
3 (n = 5334)	15.3 (8.0-22.6)	7.5 (4.5-10.4)	5.2 (3.5-6.9)	14.1 (0.0-47.9)	6.7 (4.3-9.1)
4 (Highest; n = 5955)	23.3 (14.5-32.0)	9.8 (7.4-12.2)	9.4 (6.7-12.0)	14.9 (8.8-21.1)	11.9 (9.3-14.5)
Overall (n = 22 486)	18.9 (11.6-26.2)	7.3 (5.0-9.6)	6.1 (4.1-8.1)	5.5 (0.7-10.3)	

Abbreviation: CI, confidence interval.
[a] Analysis based on random-effects pooling of Kaplan-Meier estimates from the individual studies.
[b] Categories of predicted 10-year percentage incidence of coronary heart disease, including coronary death, myocardial infarction, and angina are based on whole number cut points for scores (category 1, ≤4%; 2, 5%-7%; 3, 8%-11%; 4, ≥12%).
[c] Excludes Hoorn[21] and InCHIANTI[22] studies.

Table 8. 10-Year Total Coronary Heart Disease (CHD) Rates in Men and Women by Framingham Risk Score (FRS) Category and Ankle Brachial Index (ABI) at Baseline for All Studies Combined in the ABI Collaboration[a]

FRS Category[b]	Total		ABI							
			≤0.90		0.91-1.10		1.11-1.40		>1.40	
	No. in FRS Category	CHD, %[c]	No. in FRS Category	CHD, %[c]	No. in FRS Category	CHD, %[c]	No. in FRS Category	CHD, %[c]	No. in FRS Category	CHD, %[c]
Men										
Low (<10%)	5643	5	76	8	1076	5	4255	4	236	5
Intermediate (10%-19%)	7392	13	245	16	2069	12	4815	12	263	8
High (≥20%)	8398	23	1149	40	3406	21	3668	18	175	14
Women										
Low (<10%)	15 505	11	1083	21	6192	10	7909	9	321	11
Intermediate (10%-19%)	5563	13	558	25	2429	12	2433	11	143	13
High (≥20%)	1418	27	200	44	598	21	577	22	43	34

[a] Excludes Health in Men,[20] Hoorn,[21] and InCHIANTI[22] studies, in which nonfatal events were not available. Shaded numbers indicate individuals who would change between low (<10%), intermediate (10%-19%), and high (≥20%) risk categories from that predicted by the FRS when ABI was included. Analysis based on random-effects pooling of Kaplan-Meier estimates from the individual studies.
[b] Categories of predicted 10-year percentage incidence of coronary heart disease, including coronary death, myocardial infarction, and angina.
[c] Includes coronary death, myocardial infarction, and angina. Rates are approximate based on observed major coronary events (coronary death or myocardial infarction) adjusted by established conversion factors.[26] The number of individuals indicates those with the specified Framingham risk category and ABI level, irrespective of whether they have coronary heart disease.

cal purposes, an ABI within this range could be considered normal. Individuals with an ABI from 0.91 to 1.10 were at slightly increased risk. These results would suggest that the widely accepted high-risk cut point of 0.90 is reasonable. However, for ABI values from 0.91 to 1.10 and greater than 1.40, individuals might be advised that their risk may be slightly higher than normal levels.

The ABI Collaboration includes 16 international cohort studies. The consistency of results, especially in men (Figure 3), across a diverse spectrum of populations strengthens the validity of our findings. This consistency also was apparent despite some differences in methods of measuring the ABI and in ascertaining outcome events. We did not recalibrate the FRS, as has been suggested in populations very different from that in Framingham,[31] because in our collaboration there was no evidence that particular studies had substantially worse calibration than others and also the FRS when used in routine clinical practice is not usually calibrated to the local population. Although the area under receiver operating characteristic curves examining the added effect of the ABI are presented, from a clinical perspective, the added value of the ABI is the extent to which its inclusion reclassifies patient risk at an individual level.[32]

Other indicators of asymptomatic atherosclerosis, notably coronary artery calcium score and carotid intima media thickness have been evaluated as incremental risk predictors to the FRS. Population studies of apparently healthy individuals have suggested that coronary artery calcium score may provide added value,[33,34] particularly in discriminating high- and low-risk individuals with an intermediate FRS (predicted 10-year coronary event risk between 10% and 20%).[35] In the Atherosclerosis Risk in Communities study,[36] inclusion of carotid intima media thickness had a modest effect on the area under the receiver operating characteristic curve for the prediction of CHD using traditional risk factors. Likewise, in patients with dyslipidaemia[37] and diabetes,[38] a combination of carotid intima media thickness and FRS improved prediction compared with FRS alone. We are not aware, however, of reports of any direct comparisons in the same study of the additional values in which different measures of asymptomatic atherosclerosis (eg, coronary artery calcium vs carotid intima media thickness) make to FRS prediction in the general population.

The ABI is potentially a useful tool for prediction of cardiovascular risk. In contrast to measurement of coronary artery calcium and carotid intima media thickness, it has the advantage of ease of use in the primary care physician's office and in community settings. The equipment is inexpensive—a handheld Doppler costs less than $600. The procedure is simple, taking less than 10 to 15 minutes,[39,40] and can be performed by a suitably trained nurse or other health care professional. Technological advances to make the test quicker and easier to apply are being investigated, including automatic pressure measurement at the ankle.[41] Given the noninvasiveness of the test and minimal discomfort, patient acceptability is high. Variability is comparable with that of routine blood pressure[42,43] and individuals with borderline results may benefit from a repeated measure at a different visit.[43]

Although widely used in specialist vascular clinics, the ABI is rarely applied in routine clinical practice. Barriers to its use include: (1) most clinicians are not aware that a low ABI is a marker of cardiovascular risk; (2) it is perceived as a specialist test for use only by vascular surgeons and physicians; and (3) most clinicians would not know how to perform the test. Physician education would be essential in promoting use of the ABI in practice. Furthermore, in a survey of physicians primed to use the ABI in 1 program in the United States, time constraints, lack of reimbursement, and staff availability were barriers to use of the ABI, each reported by around half the physicians.[40]

The yield of a screening test also is important. Our results indicate that a proportion of men and women having an ABI test would be placed in a different risk category. However, this proportion may vary considerably by age because the prevalence of a low ABI is known to increase substantially with age. For example, in the United States in 2000, the prevalence of an ABI lower than 0.90 in non–Hispanic white men aged 40 to 49 years was 1.4% but was 22.6% in those aged 80 years or older.[44] Significantly higher prevalences were found in blacks. In 12 300 men free of cardiovascular disease in the general population in Scotland, the prevalence of an ABI of 0.90 or less in those aged 50 to 54 years was 3.7% but was 12.7% in those aged 75 years or older.[45] While recognizing that most risk factors also increase with age, it is likely that the added yield of a low ABI is age-related.

Recently published guidelines by the American Heart Association and the American College of Cardiology,[46] the Transatlantic Inter-Society Consensus Working Group,[47] and the Fourth Joint European Task Force[48] have suggested that the ABI should be considered for the purposes of cardiovascular risk assessment. The results of our study indicate that, when using the FRS, this may indeed be justified to improve prediction of cardiovascular risk and provision of advice on ways to reduce that risk. A new risk equation incorporating the ABI and relevant Framingham risk variables could more accurately predict risk and our intention is to develop and validate such a model in our combined data set. Cost-effectiveness modeling of the effect of using the ABI on long-term clinical outcomes also would be of interest, as has been recommended recently by an American Heart Association expert working group on screening for atherosclerotic peripheral vascular disease (Michael H. Criqui, MD, University of California San Diego, written communication, January 2008). A cost-effectiveness analysis also would be useful because successful implemen-

tation of the ABI in programs for assessment of cardiovascular risk would require a change in reimbursement regulations in some countries.

Author Contributions: Drs Fowkes and Murray had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.
Study concept and design: Fowkes, Murray, Newman, Bouter, Stehouwer, Stoffers, Hooi, Hofman, Criqui, Fronek, Hiatt, Guralnik.
Acquisition of data: Fowkes, Butcher, Heald, Chambless, Folsom, Hirsch, DeBacker, Wautrecht, Kornitzer, Cushman, Sutton-Tyrrell, D'Agostino, Murabito, Norman, Jamrozik, Curb, Masaki, Rodríguez, Dekker, Heine, Nijpels, Stehouwer, Ferrucci, McDermott, Stoffers, Hooi, Knottnerus, Ogren, Hedblad, Witteman, Breteler, Hunink, Criqui, Langer, Hiatt, Hamman, Guralnik.
Analysis and interpretation of data: Fowkes, Murray, Butcher, Heald, R. Lee, Dramaix, Kornitzer, Newman, A. Lee, Price, Murabito, Stehouwer, Ferrucci, Hooi, Knottnerus, Ogren, Witteman, Criqui, Fronek, Hiatt, Resnick.
Drafting of the manuscript: Fowkes, Murray, Newman, A. Lee.
Critical revision of the manuscript for important intellectual content: Fowkes, Murray, Butcher, Heald, R. Lee, Chambless, Folsom, Hirsch, Dramaix, DeBacker, Wautrecht, Kornitzer, Newman, Cushman, Sutton-Tyrrell, Price, D'Agostino, Murabito, Norman, Jamrozik, Curb, Masaki, Rodríguez, Dekker, Bouter, Heine, Nijpels, Stehouwer, Ferrucci, McDermott, Stoffers, Hooi, Knottnerus, Ogren, Hedblad, Witteman, Breteler, Hunink, Hofman, Criqui, Langer, Fronek, Hiatt, Hamman, Resnick, Guralnik.
Statistical analysis: Murray, Butcher, R. Lee, Dramaix, D'Agostino, Dekker, Stehouwer, Knottnerus.
Obtained funding: Fowkes, Sutton-Tyrrell, D'Agostino, Norman, Curb, Rodríguez, Dekker, Knottnerus, Breteler, Hofman, Criqui.
Administrative, technical, or material support: Fowkes, Heald, Kornitzer, Newman, Cushman, A. Lee, Masaki, Stehouwer, Hedblad, Hofman, Criqui, Guralnik.
Study supervision: Fowkes, Murray, DeBacker, Price, Dekker, Bouter, Heine, Stehouwer, Knottnerus, Ogren, Hofman, Criqui, Fronek, Guralnik.
Financial Disclosures: Drs Fowkes and McDermott reported receiving honoraria and consulting fees and Dr Norman reported receiving a research grant from Sanofi-Aventis/BMS for purposes other than for this research. Dr McDermott reported receiving consulting fees from Hutchison Technology and educational honoraria from Otsuka Pharmaceuticals. Dr Ogren is an employee of AstraZeneca's Research and Development. No other authors reported financial disclosures.
Funding/Support: Sanofi-Aventis/BMS provided an unrestricted educational grant for data processing and initial statistical analysis. The Framingham Offspring Study is supported by National Institutes of Health grant NOI-HC-25195. The InCHIANTI and Women's Health and Aging studies are supported by the Intramural Research Program, National Institute on Aging, National Institutes of Health.
Role of the Sponsor: Sanofi-Aventis/BMS 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.
Ankle Brachial Index Collaboration Authors: F. G. R. Fowkes, G. D. Murray, I. Butcher, C. L. Heald, R. J. Lee (coordinating center); L. E. Chambless, A. R. Folsom, A. T. Hirsch (Atherosclerosis Risk in Communities [ARIC] Study); M. Dramaix, G. deBacker, J-C. Wautrecht, M. Kornitzer (Belgian Physical Fitness Study); A. B. Newman, M. Cushman, K. Sutton-Tyrrell (Cardiovascular Health Study); F. G. R. Fowkes, A. J. Lee, J. F. Price (Edinburgh Artery Study); R. B. d'Agostino, J. M. Murabito (Framingham Offspring Study); P. E. Norman, K. Jamrozik (Health in Men Study); J. D. Curb, K. H. Masaki, B. L. Rodríguez (Honolulu Heart Program); J. M. Dekker, L. M. Bouter, R. J. Heine, G. Nijpels, C. D. A. Stehouwer (Hoorn Study); L. Ferrucci, M. M. McDermott (InCHIANTI Study); H. E. Stoffers, J. D. Hooi, J. A. Knottnerus (Limburg PAOD Study); M. Ogren, B. Hedblad (Men Born in 1914 Study); J. C. Witteman, M. M. B. Breteler, M. G. M. Hunink, A. Hofman (Rotterdam Study); M. H. Criqui, R. D. Langer, A. Fronek (San Diego Study); W. R. Hiatt, R. Hamman (San Luis Valley Diabetes Study); H. E. Resnick (Strong Heart Study); J. Guralnik, M. M. McDermott (Women's Health and Aging Study).

REFERENCES

1. Greenland P, Smith SC Jr, Grundy SM. Improving coronary heart disease risk assessment in asymptomatic people: role of traditional risk factors and noninvasive cardiovascular tests. *Circulation*. 2001;104(15):1863-1867.
2. Anderson KM, Odell PM, Wilson PW, Kannel WB. Cardiovascular disease risk profiles. *Am Heart J*. 1991;121(1 pt 2):293-298.
3. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation*. 1998;97(18):1837-1847.
4. Brindle P, Beswick AD, Fahey T, Ebrahim SB. Accuracy and impact of risk assessment in the primary prevention of cardiovascular disease: a systematic review. *Heart*. 2006;92(12):1752-1759.
5. Wannamethee SG, Shaper AG, Lennon L, Morris RW. Metabolic syndrome vs Framingham risk score for prediction of coronary heart disease, stroke, and type 2 diabetes mellitus. *Arch Intern Med*. 2005;165(22):2644-2650.
6. Cushman M, Arnold AM, Psaty BM, et al. C-reactive protein and the 10-year incidence of coronary heart disease in older men and women: the Cardiovascular Health Study. *Circulation*. 2005;112(1):25-31.
7. Tsimikas S, Willerson JT, Ridker PM. C-reactive protein and other emerging blood biomarkers to optimize risk stratification of vulnerable patients. *J Am Coll Cardiol*. 2006;47(8)(suppl):C19-C31.
8. Newman AB, Siscovick DS, Manolio TA, et al; Cardiovascular Heart Study (CHS) Collaborative Research Group. Ankle-arm index as a marker of atherosclerosis in the Cardiovascular Health Study. *Circulation*. 1993;88(3):837-845.
9. Weatherley BD, Nelson JJ, Heiss G, et al. 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.
10. Newman AB, Shemanski L, Manolio TA, et al; Cardiovascular Health Study Group. Ankle-arm index as a predictor of cardiovascular disease and mortality in the Cardiovascular Health Study. *Arterioscler Thromb Vasc Biol*. 1999;19(3):538-545.
11. Abbott RD, Petrovitch H, Rodriguez BL, et al. Ankle/brachial blood pressure in men >70 years of age and the risk of coronary heart disease. *Am J Cardiol*. 2000;86(3):280-284.
12. Resnick HE, Lindsay RS, McDermott MM, et al. Relationship of high and low ankle brachial index to all-cause and cardiovascular disease mortality: the Strong Heart Study. *Circulation*. 2004;109(6):733-739.
13. Leng GC, Fowkes FG, Lee AJ, Dunbar J, Housley E, Ruckley CV. Use of ankle brachial pressure index to predict cardiovascular events and death: a cohort study. *BMJ*. 1996;313(7070):1440-1444.
14. Hooi JD, Kester AD, Stoffers HE, Rinkens PE, Knottnerus JA, van Ree JW. Asymptomatic peripheral arterial occlusive disease predicted cardiovascular morbidity and mortality in a 7-year follow-up study. *J Clin Epidemiol*. 2004;57(3):294-300.
15. Ogren M, Hedblad B, Isacsson SO, Janzon L, Jungquist G, Lindell SE. Non-invasively detected carotid stenosis and ischaemic heart disease in men with leg arteriosclerosis. *Lancet*. 1993;342(8880):1138-1141.
16. van der Meer IM, Bots ML, Hofman A, del Sol AI, van der Kuip DA, Witteman JC. Predictive value of noninvasive measures of atherosclerosis for incident myocardial infarction: the Rotterdam Study. *Circulation*. 2004;109(9):1089-1094.
17. Kornitzer M, Dramaix M, Sobolski J, Degre S, De Backer G. Ankle/arm pressure index in asymptomatic middle-aged males: an independent predictor of ten-year coronary heart disease mortality. *Angiology*. 1995;46(3):211-219.
18. Whitehead A. *Meta-Analysis of Controlled Clinical Trials*. Chichester, England: John Wiley & Sons Ltd; 2002.
19. 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*. 2002;143(6):961-965.
20. Fowler B, Jamrozik K, Norman P, Allen Y. Prevalence of peripheral arterial disease: persistence of excess risk in former smokers. *Aust N Z J Public Health*. 2002;26(3):219-224.
21. Jager A, Kostense PJ, Ruhe HG, et al. Microalbuminuria and peripheral arterial disease are independent predictors of cardiovascular and all-cause mortality, especially among hypertensive subjects: five-year follow-up of the Hoorn Study. *Arterioscler Thromb Vasc Biol*. 1999;19(3):617-624.
22. 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*. 2004;52(3):405-410.
23. Criqui MH, Langer RD, Fronek A, et al. Mortality over a period of 10 years in patients with peripheral arterial disease. *N Engl J Med*. 1992;326(6):381-386.
24. Hiatt WR, Hoag S, Hamman RF. Effect of diagnostic criteria on the prevalence of peripheral arterial disease: the San Luis Valley Diabetes Study. *Circulation*. 1995;91(5):1472-1479.
25. McDermott MM, Fried L, Simonsick E, Ling S, Guralnik JM. Asymptomatic peripheral arterial disease is independently associated with impaired lower extremity functioning: the women's health and aging study. *Circulation*. 2000;101(9):1007-1012.
26. Grundy SM, Pasternak R, Greenland P, Smith S, Fuster V. Assessment of cardiovascular risk by use of multiple-risk-factor assessment equations: a statement for health care professionals from the American Heart Association and the American College of Cardiology. *J Am Coll Cardiol*. 1999;34(4):1348-1359.
27. 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(19):2486-2497.
28. O'Hare AM, Katz R, Shlipak MG, Cushman M, Newman AB. Mortality and cardiovascular risk across the ankle-arm index spectrum: results from the Cardiovascular Health Study. *Circulation*. 2006;113(3):388-393.
29. McDermott MM, Liu K, Criqui MH, et al. Ankle brachial index and subclinical cardiac and carotid disease: the Multi-Ethnic Study of Atherosclerosis. *Am J Epidemiol*. 2005;162(1):33-41.

30. Everhart JE, Pettitt DJ, Knowler WC, Rose FA, Bennett PH. Medial arterial calcification and its association with mortality and complications of diabetes. *Diabetologia*. 1988;31(1):16-23.
31. D'Agostino RB Sr, Grundy S, Sullivan LM, Wilson P; CHD Risk Prediction Group. Validation of the Framingham coronary heart disease prediction scores: results of a multiple ethnic groups investigation. *JAMA*. 2001;286(2):180-187.
32. Cook NR. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation*. 2007;115(7):928-935.
33. Greenland P, LaBree L, Azen SP, Doherty TM, Detrano RC. Coronary artery calcium score combined with Framingham score for risk prediction in asymptomatic individuals. *JAMA*. 2004;291(2):210-215.
34. Arad Y, Goodman KJ, Roth M, Newstein D, Guerci AD. Coronary calcification, coronary disease risk factors, C-reactive protein, and atherosclerotic cardiovascular disease events: the St Francis Heart Study. *J Am Coll Cardiol*. 2005;46(1):158-165.
35. Greenland P, Bonow RO, Brundage BH, et al; American College of Cardiology Foundation Clinical Expert Consensus Task Force; Society of Atherosclerosis Imaging and Prevention; Society of Cardiovascular Computed Tomography. ACCF/AHA 2007 clinical expert consensus document on coronary artery calcium scoring by computed tomography in global cardiovascular risk assessment and in evaluation of patients with chest pain: a report of the American College of Cardiology Foundation Clinical Expert Consensus Task Force. *Circulation*. 2007;115(3):402-426.
36. Chambless LE, Folsom AR, Sharrett AR, et al. Coronary heart disease risk prediction in the Atherosclerosis Risk in Communities (ARIC) study. *J Clin Epidemiol*. 2003;56(9):880-890.
37. Baldassarre D, Amato M, Pustina L, et al. Measurement of carotid artery intima-media thickness in dyslipidemic patients increases the power of traditional risk factors to predict cardiovascular events. *Atherosclerosis*. 2007;191(2):403-408.
38. Bernard S, Serusclat A, Targe F, et al. Incremental predictive value of carotid ultrasonography in the assessment of coronary risk in a cohort of asymptomatic type 2 diabetic subjects. *Diabetes Care*. 2005;28(5):1158-1162.
39. Farkouh ME, Oddone EZ, Simel DL. Improving the clinical examination for a low ankle-brachial index. *Int J Angiol*. 2002;11(1):1067-1711.
40. Mohler ER III, Treat-Jacobson D, Reilly MP, et al. Utility and barriers to performance of the ankle-brachial index in primary care practice. *Vasc Med*. 2004;9(4):253-260.
41. Jönsson B, Laurent C, Eneling M, Skau T, Lindberg L-G. Automatic ankle pressure measurements using PPG in ankle-brachial pressure index determination. *Eur J Vasc Endovasc Surg*. 2005;30(4):395-401.
42. Baker JD, Dix DE. Variability of Doppler ankle pressures with arterial occlusive disease: an evaluation of ankle index and brachial-ankle pressure gradient. *Surgery*. 1981;89(1):134-137.
43. Fowkes FGR, Housley E, Macintyre CCA, Prescott RJ, Ruckley CV. Variability of ankle and brachial systolic pressures in the measurement of atherosclerotic peripheral arterial disease. *J Epidemiol Community Health*. 1988;42(2):128-133.
44. Allison MA, Ho E, Denenberg JO, et al. Ethnic-specific prevalence of peripheral arterial disease in the United States. *Am J Prev Med*. 2007;32(4):328-333.
45. Price JF, Stewart MC, Douglas AF, Murray GD, Fowkes GF. Frequency of a low ankle brachial index in the general population by age, sex and deprivation: cross-sectional survey of 28 980 men and women. *Eur J Cardiovasc Prev Rehabil*. 2008;15(3):370-375.
46. Hirsch AT, Haskal ZJ, Hertzer NR, et al; Vascular Disease Foundation. 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(11):e463-e654.
47. Norgren L, Hiatt WR, Dormandy JA, et al. Inter-Society Consensus for the Management of Peripheral Arterial Disease (TASC II). *Eur J Vasc Endovasc Surg*. 2007;33(suppl 1):S1-S70.
48. Graham I, Atar D, Borch-Johnsen K, et al; European Society of Cardiology (ESC); European Association for Cardiovascular Prevention and Rehabilitation (EACPR); Council on Cardiovascular Nursing; European Association for Study of Diabetes (EASD); International Diabetes Federation Europe (IDF-Europe); European Stroke Initiative (EUSI); Society of Behavioural Medicine (ISBM); European Society of Hypertension (ESH); WONCA Europe (European Society of General Practice/Family Medicine); European Heart Network (EHN); European Atherosclerosis Society (EAS). European guidelines on cardiovascular disease prevention in clinical practice: full text: Fourth Joint Task Force of the European Society of Cardiology and other societies on cardiovascular disease prevention in clinical practice (constituted by representatives of nine societies and by invited experts). *Eur J Cardiovasc Prev Rehabil*. 2007;14(suppl 2):S1-S113.

PERSPECTIVES

Numeracy and Communication with Patients: They Are Counting on Us

Andrea J. Apter, MD, MA, MSc[1], Michael K. Paasche-Orlow, MD, MA, MPH[2], Janine T. Remillard, PhD[3], Ian M. Bennett, MD, PhD[4], Elana Pearl Ben-Joseph, MD[5], Rosanna M. Batista, MPH[6], James Hyde, MA, SM[7], and Rima E. Rudd, ScD[8]

[1]Division of Pulmonary, Allergy, Critical Care Medicine, Department of Medicine, University of Pennsylvania, Philadelphia, PA, USA; [2]Section of General Internal Medicine, Department of Medicine, Boston University School of Medicine, Boston, MA, USA; [3]Graduate School of Education, University of Pennsylvania, Philadelphia, PA, USA; [4]Department of Family Medicine and Community Health, University of Pennsylvania, Philadelphia, PA, USA; [5]Nemours Center for Children's Health Media, Alfred I. DuPont Hospital for Children, Wilmington, DE, USA; [6]Research Study Coordinator, Children's Hospital, Boston, MA, USA; [7]Department of Public Health and Family Medicine, Tufts University School of Medicine, Boston, MA, USA; [8]Department of Society, Human Development, and Health, Harvard School of Public Health, Boston, MA, USA
.

Patient-centered interactive communication between physicians and patients is recommended to improve the quality of medical care. Numerical concepts are important components of such exchanges and include arithmetic and use of percentages, as well as higher level tasks like estimation, probability, problem-solving, and risk assessment - the basis of preventive medicine. Difficulty with numerical concepts may impede communication. The current evidence on prevalence, measurement, and outcomes related to numeracy is presented, along with a summary of best practices for communication of numerical information. This information is integrated into a hierarchical model of mathematical concepts and skills, which can guide clinicians toward numerical communication that is easier to use with patients.

KEY WORDS: numeracy; health literacy; health communication; risk.
J Gen Intern Med 23(12):2117–24
DOI: 10.1007/s11606-008-0803-x

Received January 22, 2008
Revised August 12, 2008
Accepted September 3, 2008
Published online October 2, 2008

INTRODUCTION

Inadequate health literacy is associated with poorer communication and health outcomes, higher health costs, and likely contributes to health disparities.[1–11] This research has focused on a narrow set of literacy skills relating primarily to reading comprehension; however, limited numeracy, a component of literacy, is frequently unrecognized and limits patients' ability to communicate with health professionals. Limited numeracy skills will also hamper a patient's ability to understand health information, to make decisions related to health and health care, and may be linked to worse health outcomes.[12–15] Numeracy is increasingly relevant as the promotion of shared decision-making and the use of electronic information have increased the amount of quantitative information patients must comprehend.[16]

WHAT IS NUMERACY?

As an element of health literacy, numeracy comprises basic math skills needed for health-related activities such as timing, scheduling, and dosing of medications as well as numeric concepts needed to understand and act upon directions and recommendations given by health-care providers.[6,12,17] Numeric concepts include higher level tasks like estimation, probability, problem-solving (the ability to decipher when and how to apply numerical skills), understanding variability and error in measurement, and risk assessment. As illustrated through vignettes in Text Box 1, these skills are central to many elements of the clinical encounter.[18,19]

Text Box 1. Vignettes illustrating challenges to patient-clinician communication surrounding numerical concepts.

Vignette	Scenario	Mathematics concept
#1	A 22-year-old woman with unstable asthma is asked to record peak flow readings in the grid provided with the device. She is afraid to tell her doctors that she does not understand how to graph the numbers	Reading numbers, counting
#2	A 55-year-old man, hospitalized for a COPD exacerbation, is discharged with a bottle containing 5-mg prednisone tablets. He is told to take 30 mg in the morning for 5 days. When asked how many pills he should take tomorrow morning, he is unsure	Arithmetic operations*
#3	A mother examines the growth chart of her 6 month old. She sees the line of growth rising consistently along the 10th percentile and feels anxious that her child is too small or underweight. She (erroneously) suspects that she lacks sufficient breast milk and decides to stop nursing, despite the doctor's reassurance that her baby's growth is normal	Estimates, trends, graph reading
#4	A 50-year-old man weighs 275 lbs. His cardiologist advises him that even a 5% weight loss will greatly improve his health. The man has no idea how to determine how many pounds he should lose	Percentage, relative versus absolute values
#5	A physician prescribes alendronate for osteoporosis. The patient asks how likely it is that she will avoid a hip fracture by taking this medication. Her physician responds "the number needed to treat is 15: if 15 patients are treated, 1 will benefit."[90] The woman is confused	Probability, risk

Arithmetic operations = addition, subtraction, multiplication, division

MEASURING NUMERACY

The 1992 National Adult Literacy Survey (NALS)[20] and the 2003 National Assessment of Adult Literacy (NAAL)[21,22] included assessment of quantitative skills, the application of basic math. Findings indicate that 22% of American adults possess no more than the most simple and concrete quantitative skills. Another 33% of adults have only basic quantitative skills. These surveys evaluate numeracy in a written format that requires reading comprehension, thereby intermeshing reading and numerical activities and complicating independent assessments of numeracy. The Adult Literacy and Lifeskills Survey, the most recent international assessment of adult literacy skills, included measures of numeracy, defined as 'the ability to interpret, apply, and communicate mathematical information in commonly encountered situations."[23,24] Of the six countries participating, the US ranked fifth, below Switzerland, Norway, Bermuda, and Canada.[25]

Relatively few health-related assessment tools are dedicated to or contain quantitative items.[1,11,13,26–34] They vary in number of items, administration, and mathematical content.[35] Topics range from items assessing arithmetic skills like counting to more complicated skills like calculations of probability and risk. One unique measure tests subjective self-assessment of numeracy skills, which has been shown to correlate with actual mathematical skills.[32]

Results of numeracy tests in health settings are discouraging.[13,33,34,36] For example, only 16% of women participating in a study of basic percentage and probability concepts related to the benefits of screening for breast cancer answered all items correctly.[13] Similarly, 16% of study participants answered all items correctly in a test of common asthma self-management concepts requiring simple arithmetic and percentage computations.[1]

Most numeracy measures are self-administered, requiring reading comprehension.[13,26–34] Researchers testing understanding of food labels found that even patients with higher prose literacy had difficulty interpreting numerical information on labels.[37] Similarly, even though scores on the Asthma Numeracy Questionnaire (ANQ) generally correlated with scores of the Short Test of Functional Health Literacy in Adults (STOFLHA) ($r=0.34$, $p=0.004$), individuals who scored well on the reading test did not necessarily score well on the ANQ.[1] Researchers, using tests containing probability, positive predictive value, and other complex concepts, found that better numeracy was associated with more education, being male, and being white.[7,11,34] At the same time, a study of women's capacity to estimate the chance of breast cancer survival and benefit of screening mammography found that black compared with white women were more likely to make an accurate assessment of cancer survival and women who did not graduate from high school were more accurate in assessing mammography benefit.[38]

Even well-educated patients may have trouble converting proportions to percent or understanding simple probability and risk.[13,19,32,39–43] Among medical students attending a seminar on risk-communication, 23% got at least one item wrong on a numeracy test assessing risk.[13] Although 90% were able to determine which of two drugs offered greater benefit when information was presented in terms of relative risk reduction, absolute risk reduction, number-needed-to-treat or a combination of these concepts, only 61% could calculate how much one of the drugs reduced disease risk.[43] In another study, almost half of the doctors surveyed made different treatment recommendations when identical data were presented in a relative versus an absolute risk format.[44]

NUMERACY AND HEALTH OUTCOMES

The literature in health literacy research offers substantial evidence of links between literacy skills and health outcomes.[16] However, research on the relationship between numeracy and health is scant. Limited numeracy has been associated with poorer anti-coagulation control[45] and poorer diabetes self-management.[11] Additionally, limited numeracy was associated with a history of more hospitalizations and ED visits for adults with asthma.[1] However, one study found that

correct scores on probability questions were not associated with being up-to-date with colorectal screening or mammography,[46] and another reported that understanding numerical concepts in nutrition labels was not associated with blood pressure or cholesterol levels.[47]

CONCEPTUAL MODEL FOR COMMUNICATION OF NUMERICAL INFORMATION

Knowing that the average numeracy skills of US adults are limited, clinicians may likely struggle with the communication of critical health information. In Table 1 we introduce a conceptual model of elements necessary for communication of health information with numeric concepts. The model is based on Golbeck's four overlapping categories of numerical information: basic (e.g., ability to identify and read numbers), computational (e.g., counting and arithmetic), analytical (e.g., inference, estimation, proportion, percentage, frequencies, basic graphs), and statistical (e.g., basic probability, statistics, and risk assessment).[48] It presents numerical concepts arranged by difficulty as assessed and taught by educators.[49] We selected those concepts frequently appearing in health communication.[19] Related concepts are adjacent and similarly colored.[48–50] Relatively easy tasks like reading or locating numbers, such as those on a peak flow meter, fall within the

Table 1. Numeracy and Patient-Physician Communication: A Hierarchy of Numerical Complexity and Comprehension

Numeracy element*	Level of Patient Mastery Required (demand on patient)		
	Describe	Interpret	Decision-Making
Reading numbers, counting, telling time	V1¶		
Arithmetic operations			V2
Estimation of size, trend	V1	V3	
Frequency	V5		
Percentage		V3, V4	V4
Problem-solving† & inferring the mathematical concepts to be applied			V1, V4
Logic‡			
Reading tables	V1		
Reading graphs	V1	V3	
Reading maps			
Estimation of error, uncertainty, variability§		V3	
Relative versus absolute‖			
Risk (cumulative, relative, conditional)			V5

We propose this matrix as a conceptual model that offers a theoretical guide for communicating numerical information. It is also a framework for formulating research to improve communication of numerical information. The left column displays numerical concepts frequently used in health care, grouped by approximate level of difficulty. From left to right, the columns represent the level of comprehension patients need to perform health-related tasks. We hypothesize that patient autonomy and shared decision-making can be improved by, whenever possible, replacing a communication in one cell by one that is higher and to the left. We link this matrix to the vignettes to show where in this matrix common self-care activities and patient-clinician communication are situated.

**Numerical tasks are displayed vertically in order of difficulty with colors indicating related numerical tasks of approximately similar level of difficulty. We emphasize that the ordering of difficulty is only approximate. The categories are roughly as taught sequentially in schools and as described by educators.[49]*

†*Problem-solving is the ability to decide which numerical or logical concepts to employ in order to find a solution.*

‡*Logic: the understanding of logical operators such as and, or, not.*

§*Estimation of error/uncertainty, variability: e.g. understanding measurement differences, such as glucose of 101, 99, 102 do not indicate significant clinical differences in blood sugar; or that with a weight of 220 lbs one day and 230 lbs the next day, such large variability indicates a probable error in measurement.*

‖ *Relative versus absolute indicates the need to be able to understand and compare absolute and relative changes, particularly when absolute values are small. In V4, the patient must comprehend relative compared with absolute weight.*

¶*V1 = Vignette #1*

top left cells. More complicated tasks like estimating size and problem-solving (determining the appropriate mathematical concept to employ) are further down.

Low on Text Box 1 and Table 1 is the numerical concept most frequently studied in health communication: risk, the probability of a bad outcome.[39,48,51,52] Indeed, all preventive care revolves around risk reduction.[40,50,52] Concepts like the standard gamble, time trade-off, and number-needed-to-treat were developed to facilitate understanding of risk, but are themselves difficult to understand.[39,40,42,43,53–55] Changes in risk are sometimes presented as relative values without absolute quantities. This may be particularly hard to understand when the absolute quantities are small. For example, a patient advised to take a medication to reduce the chance of a bad outcome by 50% may not understand that the absolute risk is only 0.4% per lifetime.

In Table 1, the hierarchy of numerical concepts is depicted vertically, and the depth of comprehension needed by patients to apply these skills to describe, interpret, or make decisions is depicted horizontally. It is similar to a model of assimilation and synthesis of medical information used in medical student education.[56] We hypothesize that patient autonomy and shared decision-making can be improved by replacing a communication in one cell by one that is higher and to the left.

TECHNIQUES TO IMPROVE COMPREHENSION OF QUANTITATIVE INFORMATION

Recommended techniques fall into six groupings and include: simplification, clear formatting, omission of distracting information, appropriate framing, use of visuals, and confirmation.

Simplify the Numerical Concept

Simplifying means explaining the concept by moving as far as possible upward and to the left in the model presented in Table 1. For example, in Vignette 4, recommending that the patient lose 14 lbs rather than 5% of his weight replaces a numeracy task of deriving a percentage with the easier task of reading/telling numbers. In addition, a statement that "even a 5% weight loss will improve health" leaves the patient in the second column of interpreting rather than in the first column of describing. Such simplification does not negate the obligation to provide more detailed and complete information for those who want it.

Format for Clarity

Use of white space and simple prose captions to accompany the numerical message is recommended for all documents to ease reading. Studies indicate that employing larger rather than the smaller numbers to represent more desirable scores facilitates understanding. Thus, for a scale of 1 to 5, 5 should represent the desirable score rather than 1.[7,57] Furthermore, tables and graphs should present the most important information first and the least important last to highlight key information and to improve understanding.[7]

Remove Nonessential Information

The presence of distracting information makes any text more difficult to use.[7,58] The key is to understand what information is critical and what is extraneous.

Frame Effectively

Framing describes the packaging or presentation of information and influences its interpretation.[6,52] Patients tend to underestimate common risks and overestimate rare risks.[52] The depiction of risk as 1 in 10 may be understood differently from a risk of 10 in 100.[51] Furthermore, patients' interpretation of risk tends to be biased toward an outcome presented in a positive versus negative light, i.e., risk of chemotherapy posed as the probability of living rather than of dying.[6,59–62] A patient may not understand the significance of a cholesterol level of 160 mg/dl, until told that 160 is within the normal range of 112–200 mg/dl. Even well-educated clinicians may be influenced by framing.[52] Forensic psychologists and psychiatrists were less likely to discharge a patient when told *20 out of every 100* similar patients were estimated to commit an act of violence, compared with being told that such patients have a *20% chance* of committing an act of violence.[63] When numerical information, such as risk, is unfamiliar, reliance on framing increases and will increasingly determine which information is used in making decisions.

For communicating risk and probability, numbers rather than words are associated with a more accurate perception. Words without numbers like the words few, some, and many do not have precise meaning.[62,64] However, interpretive framing that uses explanatory phrases along with numbers can enhance communication and increase trust in the physician and belief in the health information.[65–67]

Another consideration relates to the value of framing issues within a time period. The time frame used influences risk perception.[51,68] For example, one study found that older women preferred a 10-year time frame and younger women resonating more with a 1-year time frame.[51] Framing, if used wisely, provides context and supports recipients in finding meaning in the numerical message.[52,62,67]

Use Visuals

Visuals, including tables, graphs, formatted boxes in essays, and pictures, enhance understanding.[69,70] The choice of image influences interpretation of numerical concepts and, thus, must be tailored to the patient, the numerical concept, and the health message.[69,70] For example, in Figure 1 the identical lifetime risk of breast cancer for a 50-year-old woman was displayed in different formats and presented to focus groups of women.[51,89] Figure 1a, a frequency graph with a clear numerator and denominator, was considered more understandable than the bar graph of Figure 1b, which does not provide a denominator. Denominators of 10 or 100 were easiest for focus group participants to understand, compared with larger denominators. Additionally, women tended to perceive larger risk for identical proportions if a smaller denominator was used, i.e., a 1/10 frequency graph was estimated to depict greater risk than 10 out of 100 or 100 out of 1,000. In Figure 1a human figures, an icon array,[69] were used to personalize information for women, although the focus

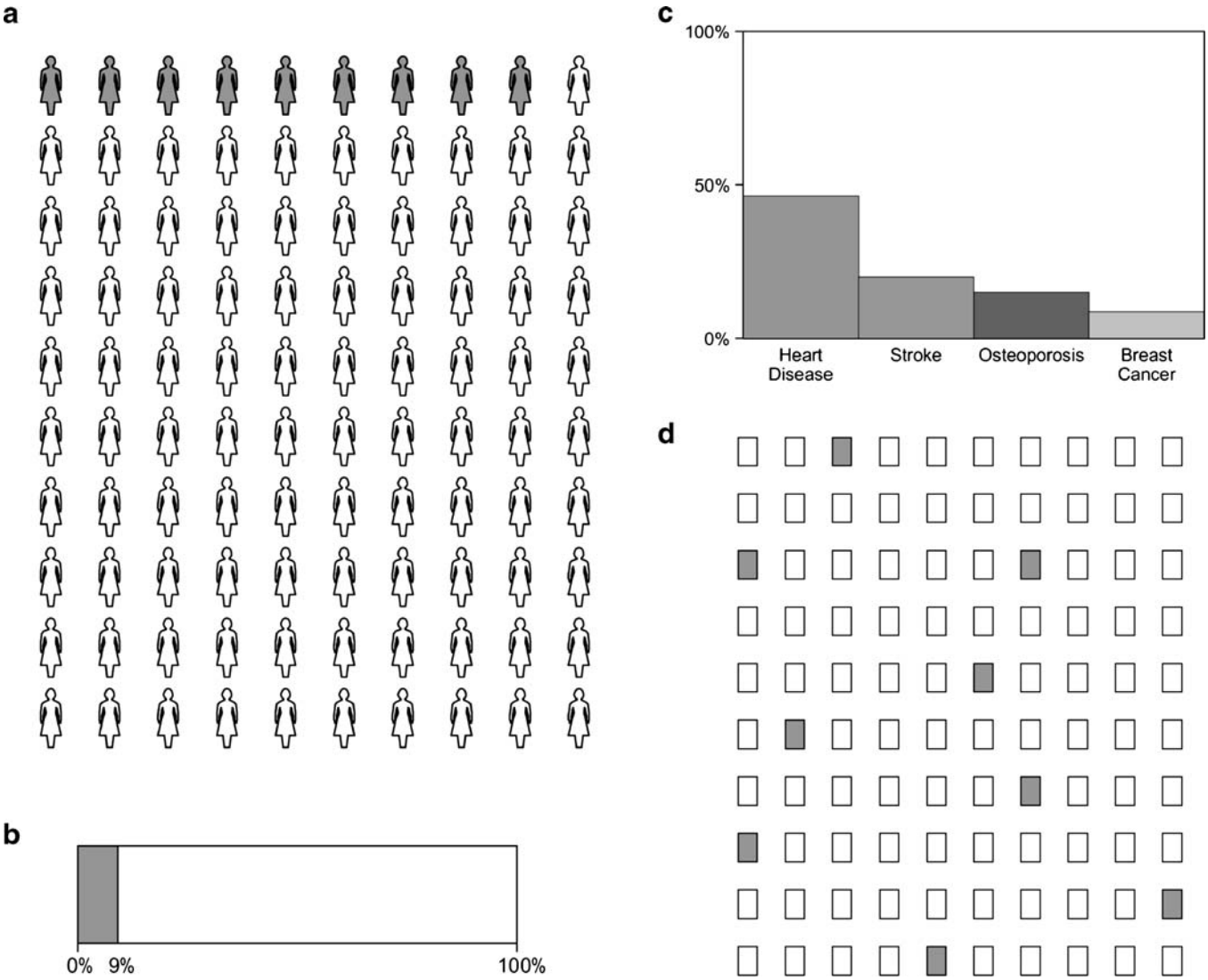


Figure 1. Examples of use of figures to convey the lifetime risk of breast cancer for a 50-year-old woman. (a) Risk is displayed as a frequency with a clear numerator and denominator. (b) This bar graph has no definite denominator; risk is displayed as a proportion rather than a frequency. (c) Multiple bar graphs depicting other comorbidities illustrates that bar graphs are an excellent format for making comparisons. (d) The random highlighting of the matrix makes it difficult to appreciate the numerator, but displays the idea of chance well. From Schapira MM, Nattinger AB, McAuliffe TL, The Influence of Graphic Format on Breast Cancer Risk Communication, *Journal of Health Communication* 2006;11:569–582, reprinted by permission of the publisher (Taylor & Francis, http://www.informaworld.com).[89]

group participants did not find the icon enhanced personal applicability.[51] The multiple bar graphs of Figure 1c depicting other comorbidities illustrates that bar graphs are an excellent format for making comparisons and depicting relative risk.[57] In Figure 1d the random highlighting was considered difficult to understand compared with consecutive highlighting (Figure 1a).[51] However, random arrays (Figure 1d) were useful in understanding chance in genetic counseling.[69]

The type of graph used can be determined by the data to be presented.[69–73] For example, part-to-whole concepts such as percentages may be emphasized with histograms and pie charts. Such formats display the denominator and may also convey relative versus absolute comparisons.[70] Line graphs are effective for communicating trends. Such trends can be distorted if the vertical scale is not representative of the true scale. Scatter plots effectively display variability.[70,74] The addition of brief captions and reference points enhances a graph's message.[52,62,70] Graphs also can distort, for example, when the numerator is displayed without the denominator.[69,75]

Pictographs have been found to improve attention and recall when they are closely linked to spoken directions or text,[71,72,76–78] and statistics presented as pictographs have been shown to reduce reliance on anecdotes and framing.[79] A study of an educational intervention to improve self-management of heart failure randomized patients to receive picture-based materials, a digital scale, and telephone follow-up. These patients had a lower rate of hospitalization or death than did those randomized to a general heart failure education brochure and usual care, and the effect was larger for patients

Table 2. Summary of Recommendations from the Literature for Presenting Numerical Concepts

Recommendation*
Use the fewest and simplest mathematical constructs[6,7,51,66] (e.g., highest and to the left on Table 1)
Remove nonessential information[7]
Order information from most to least important or along a discernible hierarchy[7]
Use several formats for presentation, e.g., verbal, quantitative, visual[7,52,57,66,67,71,77,79,91]
Consider using constructive framing or anecdotes[6,59–61,65–67,79,92]
Present benefits and risks, loss and gains, negative and positive[6,59–62]
Realize positive is more likely to be chosen[6,61]
Consider the best time frame for presenting risk[51,68]
When using graphs, use most appropriate format and explain it to the patient[52,57,69,70,72,73,79]
When applicable, show full denominator or full range of scale and explain both the numerator and denominator[51,69,70]
Tailor information to the patient[66,84]
Make communication interactive[46,66,77]
Reinforce important messages with repeat instruction[66]
Confirm comprehension[84]

*Superscripts indicate references

with low literacy.[80] Videos, interactive computer interfaces, and use of the internet all hold promise for use in patients with low literacy.[81]

Confirm Comprehension

A brief individualized assessment of numerical skill may be useful for tailoring teaching in the clinical setting, but clinical screening may be threatening to vulnerable patients. Furthermore, national studies such as the NALS and NAAL indicate that a majority of adults have limited quantitative skills. Consequently, we do not recommend clinical numeracy screening until it has been proven to benefit patients.[82] Instead, since all patients will benefit from simple explanations, we recommend that clinicians apply universal precautions (Table 2).[83] Furthermore, we should confirm comprehension of important numeric concepts with techniques such as the teach-back method.[84] This approach, asking a patient to state what will be done or what he or she will tell a family member, can be part of closing the encounter. Teach-back and other techniques noted above are helpful to all patients and particularly for those with limited numeracy.[7]

DISCUSSION

Limited numeracy is prevalent and may likely influence clinical outcomes. Increased awareness and training to help clinicians communicate successfully are important goals. To start, clinicians can use Tables 1 and 2 to guide simplification of their numerical communication.

National studies indicate that the gap in mathematics achievement between whites and blacks and Hispanics is even worse than in reading.[85,86] It is possible that health disparities in chronic disease management and for participation in screening are driven in part by poor education, particularly in mathematics. At the same time, while the concern for numeracy development is intensified for low-income populations, this concern must influence encounters with patients from middle and high income communities as well. Findings indicate that a majority of US adults do not have adequate numeracy and that K-12 mathematical instruction in the US does not prepare students for needed reasoning and problem-solving tasks.[87,88]

Focused clinical research is needed to better define the numerical concepts necessary for communicating health information and to delineate the best ways to measure and improve numeric communication. Conceptual models elucidating the pathways by which numeracy may be linked to health outcomes are needed to motivate further study. Table 1 can be considered a first approximation of such a model. The communication of numerical concepts must be studied in health-care settings, not simply in test-taking venues. It will be especially important to study patients from vulnerable populations to understand how removing unneeded complexity and improving communication around numerical concepts can decrease health disparities.[84]

Acknowledgments: *We gratefully acknowledge the suggestions and review of the Health Literacy/Health Communication Working Group of the Harvard School of Public Health.*

Acknowledgement of Support: Dr. Apter: NIH K02HL088469, R01HL073932. Dr. Bennett: NIH K23HD048915. Dr. Hyde: NIH R03 HD0540432.

Conflict of Interest: *None disclosed.*

Corresponding Author: *Andrea J. Apter, MD, MA, MSc; Division of Pulmonary, Allergy, Critical Care Medicine, Department of Medicine, University of Pennsylvania, 829 Gates Building, 3600 Spruce Street, Philadelphia, PA 19104, USA (e-mail: apter@mail.med.upenn.edu).*

REFERENCES

1. **Apter AJ, Cheng J, Small D, et al.** Asthma Numeracy Skill and Health Literacy. J Asthma. 2006;43:705–10.
2. **Baker DW, Gazmararian JA, Williams MV, et al.** Functional health literacy and the risk of hospital admission among Medicare managed care enrollees. Am J Public Health. 2002;92(8):1278–83.
3. **Gazmararian JA, Baker DW, Williams MV, et al.** Health literacy among Medicare enrollees in a managed care organization. JAMA. 1999; 281 (6):545–51.
4. **Nielson-Bohlman L, Panzer A, Kindig D, eds.** Health Literacy: A Prescription to End Confusion. Washington, D.C.: National Academies Press: 2004.
5. **Schillinger D, Grumbach K, Piette J, et al.** Association of health literacy with diabetes outcomes. JAMA. 2002;288(4):475–82.
6. **Peters E, Hibbard J, Slovic P, Dieckmann N.** Numeracy skill and the communication, comprehension, and use of risk-benefit information. Health Aff (Project Hope). 2007;26(3):741–8.
7. **Peters E, Dieckmann N, Dixon A, Hibbard JH, Mertz CK.** Less is more in presenting quality information to consumers. Med Care Res Rev. 2007;64(2):169–90.
8. **Wong MD, Shapiro MF, Boscardin WJ, Ettner SL.** Contribution of major diseases to disparities in mortality. N Engl J Med. 2002;347 (20):1585–92.
9. **Rothman RL, DeWalt DA, Malone R, et al.** Influence of patient literacy on the effectiveness of a primary care-based diabetes disease management program. JAMA. 2004;292(14):1711–6.
10. **Osborn CY, Paasche-Orlow MK, Davis TC, Wolf MS.** Health literacy an overlooked factor in understanding HIV health disparities. Am J Prev Med. 2007;33(5):374–8.
11. **Cavanaugh K, Huizinga MM, Wallston KA, et al.** Association of numeracy and diabetes control. Ann Intern Med. 2008;148(10):737–46.
12. **Montori VM, Rothman RL.** Weakness in numbers. The challenge of numeracy in health care. J Gen Intern Med. 2005;20(11):1071–2.

13. **Schwartz LM, Woloshin S, Black WC, Welch HG.** The role of numeracy in understanding the benefit of screening mammography. Ann Intern Med. 1997;127(11):966–72.
14. **Parikh NS, Parker RM, Nurss JR, Baker DW, Williams MV.** Shame and health literacy: the unspoken connection. Patient Educ Couns. 1996;27 (1):33–9.
15. **Kelly KM, Graves KD, Harper FW, Schmidt JE, Dickinson SL, Andrykowski MA.** Assessing perceptions of cancer risk: does mode of assessment or numeracy matter? Cancer Detect Prev. 2007;31(6):465–73.
16. **Rudd RE, Anderson JE, Oppenheimer S, Nath C.** Health literacy: an update of medical and public health literature. In: Comings J, Garner B, Smith C, eds. Review of Adult Learning and Literacy, Volume 7. Mathway, N.J.: Lawrence Erlbaum Associates; 2007:175–204.
17. **Gal I.** Adult Numeracy Development. Cresskill, NJ: Hampton Press; 2000.
18. **Foerch JS.** Characteristics of adult learners of mathematics. In: Gal I, ed. Adult Numeracy Development. Cresskill, NJ: Hampton Press, Inc; 2000.
19. **Ancker JS, Kaufman D.** Rethinking health numeracy: a multidisciplinary literature review. J Am Med Inform Assoc. 2007;14(6):713–21.
20. **Kirsch IS, Jungeblut A, Jenkins L, Kolstad A.** Adult literacy in America: a first look at the results of the National Adult Literacy Survey (NALS). Washington, D.C.: National Center for Education Statistics, US Department of Education; 1993.
21. National Assessment of Adult Literacy. National Center for Education Statistics;Institute of Education Sciences; US Department of Education, 2003. (Accessed 10/30/2007, 2007, at http://nces.ed.gov/naal/.)
22. **Kutner M, Greenberg E, Jin Y, Paulsen C.** The health literacy of America's adults, Results from the 2003 National Assessment of Adult Literacy. In: National Center for Education Statistics, US Department of Education; 2006.
23. **Murray TS, Clermont Y, Binkely M.** Measuring adult literacy and life skills: new frameworks for assessment. In. Ottawa, Ontario, Canada: Statistics Canada 2005:1–408.
24. **Lemke M, Miller D, Johnston J, et al.** 2003 International Adult Literacy and Lifeskills Survey (ALL). Washington, DC: National Center for Education Statistics, US Department of Education; 2005.
25. Adult Literacy and Lifeskills Survey (ALL); Highlights from the 2003 International Adult Literacy and Lifeskills Survey (ALL). National Center for Education Statistics. (Accessed 7/22/2008, at http://nces.ed.gov/surveys/ALL/issuebrief.asp.)
26. Anon. Wide Range Achievement Test (WRAT3). Wilmington, DE: Wide Range, Inc.; 1993.
27. Wide Range Achievement Test 4 (WRAT4). Wide Range, Inc. (Accessed August 11, 2008, at http://www3.parinc.com/products/product.aspx?Productid = WRAT4.)
28. **Davis TC, Long SW, Jackson RH, et al.** Rapid estimate of adult literacy in medicine: a shortened screening instrument. Fam Med. 1993;25 (6):391–5.
29. **Parker RM, Baker DW, Williams MV, Nurss JR.** The test of functional health literacy in adults: a new instrument for measuring patients' literacy skills. J Gen Intern Med. 1995;10(10):537–41.
30. **Baker DW, Williams MV, Parker RM, Gazmararian JA, Nurss J.** Development of a brief test to measure functional health literacy. Patient Educ Couns. 1999;38(1):33–42.
31. **Weiss BD, Mays MZ, Martz W, et al.** Quick assessment of literacy in primary care: the newest vital sign. Ann Fam Med. 2005;3 (6):514–22.
32. **Fagerlin A, Zikmund-Fisher BJ, Ubel PA, Jankovic A, Derry HA, Smith DM.** Measuring numeracy without a math test: development of the subjective numeracy scale. Med Decis Making. 2007;27 (5):672–80.
33. **Schwartz LM, Woloshin S, Welch HG.** Can patients interpret health information? An assessment of the medical data interpretation test. Med Decis Making. 2005;25(3):290–300.
34. **Lipkus IM, Samsa G, Rimer BK.** General performance on a numeracy scale among highly educated samples. Med Decis Making. 2001;21 (1):37–44.
35. **Donelle L, Hoffman-Goetz L, Arocha JF.** Assessing health numeracy among community-dwelling older adults. J Health Commun. 2007;12 (7):651–65.
36. **Woloshin S, Schwartz LM, Welch HG.** The effectiveness of a primer to help people understand risk: two randomized trials in distinct populations. Ann Intern Med. 2007;146(4):256–65.
37. **Rothman RL, Housam R, Weiss H, et al.** Patient understanding of food labels: the role of literacy and numeracy. Am J Prev Med. 2006;31 (5):391–8.
38. **Haggstrom DA, Schapira MM.** Black-white differences in risk perceptions of breast cancer survival and screening mammography benefit. J Gen Intern Med. 2006;21(4):371–7.
39. **Sheridan SL, Pignone MP, Lewis CL.** A randomized comparison of patients' understanding of number needed to treat and other common risk reduction formats. J Gen Intern Med. 2003;18(11):884–92.
40. **Halvorsen PA, Selmer R, Kristiansen IS.** Different ways to describe the benefits of risk-reducing treatments: a randomized trial. Ann Intern Med. 2007;146(12):848–56.
41. **Krishnan JA, Riekert KA, McCoy JV, et al.** Corticosteroid use after hospital discharge among high-risk adults with asthma. Am J Respir Crit Care Med. 2004;170(12):1281–5.
42. **Kristiansen IS, Gyrd-Hansen D, Nexoe J, Nielsen JB.** Number needed to treat: easily understood and intuitively meaningful? Theoretical considerations and a randomized trial. J Clin Epidemiol. 2002;55 (9):888–92.
43. **Sheridan SL, Pignone M.** Numeracy and the medical student's ability to interpret data. Eff Clin Pract. 2002;5(1):35–40.
44. **Forrow L, Taylor WC, Arnold RM.** Absolutely relative: how research results are summarized can affect treatment decisions. Am J Med. 1992;92(2):121–4.
45. **Estrada CA, Martin-Hryniewicz M, Peek BT, Collins C, Byrd JC.** Literacy and numeracy skills and anticoagulation control. Am J Med Sci. 2004;328(2):88–93.
46. **Aggarwal A, Speckman JL, Paasche-Orlow MK, Roloff KS, Battaglia TA.** The role of numeracy on cancer screening among urban women. Am J Health Behav. 2007;31Suppl 1S57–68.
47. **Osborn CY, Weiss BD, Davis TC, et al.** Measuring adult literacy in health care: performance of the newest vital sign. Am J Health Behav. 2007;31Suppl 1S36–46.
48. **Golbeck AL, Ahlers-Schmidt CR, Paschal AM, Dismuke SE.** A definition and operational framework for health numeracy. Am J Prev Med. 2005;29(4):375–6.
49. **Ginsburg L, Manly M, Schmitt MJ.** The Components of Numeracy. NCSALL Occasional Paper. National Center for the Study of Adult Learning and Literacy (NCSALL) 2006.
50. **Gal I.** Adults' statistical literacy: meanings, components, responsibilities. Int Stat Rev. 2002;70:1–51.
51. **Schapira MM, Nattinger AB, McHorney CA.** Frequency or probability? A qualitative study of risk communication formats used in health care. Med Decis Making. 2001;21(6):459–67.
52. **Moore RA, Derry S, McQuay HJ, Paling J.** What do we know about communicating risk? A brief review and suggestion for contextualising serious, but rare, risk, and the example of cox-2 selective and non-selective NSAIDs. Arthritis Research & Therapy. 2008;10(1):R20.
53. **Woloshin S, Schwartz LM, Moncur M, Gabriel S, Tosteson AN.** Assessing values for health: numeracy matters. Med Decis Making. 2001;21(5):382–90.
54. **Fagerlin A, Zikmund-Fisher BJ, Ubel PA, Jankovic A, Derry HA, Smith DM.** Measuring numeracy without a math test: development of the Subjective Numeracy Scale. Med Decis Making. 2007;27(5):672–80.
55. **Gyrd-Hansen D, Kristiansen IS, Nexoe J, Nielsen JB.** How do individuals apply risk information when choosing among health care interventions? Risk Anal. 2003;23(4):697–704.
56. **Pangaro L.** A new vocabulary and other innovations for improving descriptive in-training evaluations. Acad Med. 1999;74(11):1203–7.
57. **Feldman-Stewart D, Kocovski N, McConnell BA, Brundage MD, Mackillop WJ.** Perception of quantitative information for treatment decisions. Med Decis Making. 2000;20(2):228–38.
58. **Kirsch I.** The framework used in developing and interpreting the International Adult Literacy Survey. Eur J Psychol Edu. 2001;16 (3):335–61.
59. **McNeil BJ, Pauker SG, Sox HC Jr., Tversky A.** On the elicitation of preferences for alternative therapies. N Engl J Med. 1982;306(21):1259–62.
60. **Levin IP, Gaeth GJ.** How consumers are affected by the framing of attribute information before and after conuming the product. J Consume Res. 1988;15:374–8.
61. **Armstrong K, Schwartz JS, Fitzgerald G, Putt M, Ubel PA.** Effect of framing as gain versus loss on understanding and hypothetical treatment choices: survival and mortality curves. Med Decis Making. 2002;22 (1):76–83.

62. **Edwards A, Elwyn G, Mulley A.** Explaining risks: turning numerical data into meaningful pictures. BMJ (Clinical Research Ed). 2002;324 (7341):1827–30.
63. **Slovic P, Monahan J, MacGregor DG.** Violence risk assessment and risk communication: the effects of using actual cases, providing instruction, and employing probability versus frequency formats. Law Human Behav. 2000;24(3):271–96.
64. **Maat HP.** What authors and readers do with side effect information on drugs. In: Lentz L, Maat HP, eds. Discourse Analysis and Evaluation: Functional Approaches. Amsterdam/Atlanta: Ultrecht Studies in Language and Communication; 1997:111–38.
65. **Knapp P, Raynor DK, Berry DC.** Comparison of two methods of presenting risk information to patients about the side effects of medicines. Q Safe Health Care. 2004;13(3):176–80.
66. **Trevena LJ, Davey HM, Barratt A, Butow P, Caldwell P.** A systematic review on communicating with patients about evidence. J Eval Clin Pract. 2006;12(1)13–23.
67. **Gurmankin AD, Baron J, Armstrong K.** The effect of numerical statements of risk on trust and comfort with hypothetical physician risk communication. Med Decis Making. 2004;24(3):265–71.
68. **Schapira MM, Davids SL, McAuliffe TL, Nattinger AB.** Agreement between scales in the measurement of breast cancer risk perceptions. Risk Anal. 2004;24(3)665–73.
69. **Ancker JS, Senathirajah Y, Kukafka R, Starren JB.** Design features of graphs in health risk communication: a systematic review. J Am Med Inform Assoc. 2006;13(6):608–18.
70. **Lipkus IM, Hollands JG.** The visual communication of risk. J Natl Cancer Inst. 1999(25):149–63.
71. **Pylar J, Wills CE, Lillie J, Rovner DR, Kelly-Blake K, Holmes-Rovner M.** Men's interpretations of graphical information in a videotape decision aid. Health Expect. 2007;10(2):184–93.
72. **Price M, Cameron R, Butow P.** Communicating risk information: the influence of graphical display format on quantitative information perception-accuracy, comprehension and preferences. Patient Educ Couns. 2007;69:121–8.
73. **Muscatello DJ, Searles A, MacDonald R, Jorm L.** Communicating population health statistics through graphs: a randomised controlled trial of graph design interventions. BMC Med. 2006;4:33.
74. **Legge GE, Gu YC, Luebker A.** Efficiency of graphical perception. Percept Psychophys. 1989;46(4):365–74.
75. **Cleveland WS, McGill R.** Graphical perception and graphical methods for analyzing scientific data. Science. 1985;229:828–33.
76. **Houts PS, Bachrach R, Witmer JT, Tringali CA, Bucher JA, Localio RA.** Using pictographs to enhance recall of spoken medical instructions. Patient Educ Couns. 1998;35(2):83–8.
77. **Houts PS, Doak CC, Doak LG, Loscalzo MJ.** The role of pictures in improving health communication: a review of research on attention, comprehension, recall, and adherence. Patient Educ Couns. 2006;61 (2):173–90.
78. **Houts PS, Witmer JT, Egeth HE, Loscalzo MJ, Zabora JR.** Using pictographs to enhance recall of spoken medical instructions II. Patient Educ Couns. 2001;43(3):231–42.
79. **Fagerlin A, Wang C, Ubel PA.** Reducing the influence of anecdotal reasoning on people's health care decisions: is a picture worth a thousand statistics? Med Decis Making. 2005;25(4):398–405.
80. **Dewalt DA, Malone RM, Bryant ME, et al.** A heart failure self-management program for patients of all literacy levels: A randomized, controlled trial [ISRCTN11535170]. BMC Health Serv Res. 2006;6(1):30.
81. **Jibaja-Weiss ML, Volk RJ.** Utilizing computerized entertainment education in the development of decision AIDS for lower literate and naive computer users. J Health Commun. 2007;12(7):681–97.
82. **Paasche-Orlow MK, Wolf MS.** Evidence does not support clinical screening of literacy. J Gen Intern Med. 2008;23(1):100–2.
83. **Paasche-Orlow MK, Schillinger D, Greene SM, Wagner EH.** How health care systems can begin to address the challenge of limited literacy. J Gen Intern Med. 2006;21(8):884–7.
84. **Paasche-Orlow MK, Riekert KA, Bilderback A, et al.** Tailored education may reduce health literacy disparities in asthma self-management. Am J Respir Crit Care Med. 2005;172(8):980–6.
85. Grade Report Card. National Center for Education Statistics, US Department of Education, 2007. http://nces.ed.gov/nationsreportcard/pdf/main2007/2007494_3.pdf (Accessed 8/11/2008).
86. **Perle M, Moran R.** NAEP 2004 Trends in Academic Progress: Three Decades of Student Performance in Reading and Mathematics (NCES 2005–464). In. Washington, DC: US Department of Education, Institute of Education Sciences, National Center for Education Statistics; 2005.
87. **Schmidt WH, McKnight CC, Raizen SA.** A splintered vision: an investigation of US science and mathematics education. Dordrecth, The Netherlands: Kluwer Academic Publishers; 1997.
88. **Kilpatrick J, Swafford J, Findell B.** Adding it Up: Helping Children Learn Mathematics. Washington, D.C.: Mathematics Learning Study Committee, Center for Education, Division of Behavioral and Social Sciences and Education, National Research Council, National Academy Press; 2001.
89. **Schapira MM, Nattinger AB, McAuliffe TL.** The influence of graphic format on breast cancer risk communication. J Health Commun. 2006;11(6):569–82.
90. **Cummings SR, Black DM, Thompson DE, et al.** Effect of alendronate on risk of fracture in women with low bone density but without vertebral fractures: results from the Fracture Intervention Trial. JAMA. 1998;280 (24):2077–82.
91. **DeWalt DA, Malone RM, Bryant ME, et al.** A heart failure self-management program for patients of all literacy levels: a randomized, controlled trial [ISRCTN11535170]. BMC Health Serv Res. 2006;6:30.
92. **Gurmankin AD, Baron J, Hershey JC, Ubel PA.** The role of physicians' recommendations in medical treatment decisions. Med Decis Making. 2002;22(3):262–71.

EDITORIAL

Genetic Testing for Warfarin Dosing? Not Yet Ready for Prime Time

Henry I. Bussey, Pharm.D., FCCP, FAHA, Ann K. Wittkowsky, Pharm.D., FASHP, FCCP, Elaine M. Hylek, M.D., M.P.H., and Marie B. Walker, B.B.A.

Key Words: genetic testing, warfarin, dosage.
(Pharmacotherapy 2008;28(2):141–143)

Recently, much media attention has been focused on genetic testing as a means to determine the appropriate warfarin dosage for an individual patient, thereby substantially reducing the risk of bleeding or clotting events.[1–4] At the Anticoagulation Forum National Conference on Anticoagulation Therapy held in Chicago in May 2007, genetic testing for warfarin dosing was a hot topic. Even the United States Food and Drug Administration (FDA) has weighed in on this issue.[5] In addition, the American Enterprise Institute–Brookings Joint Center (with input from the FDA) has published a document that reached some very impressive conclusions.[6] Specifically, the report concludes, "We estimate that formally integrating genetic testing into routine warfarin therapy could allow American warfarin users to avoid 85,000 serious bleeding events and 17,000 strokes annually. We estimate the reduced health care spending from integrating genetic testing into warfarin therapy to be $1.1 billion annually, with a range of about $100 million to $2 billion."

So, what's the story? Can genetic testing be used to determine the right warfarin dosage? Does such an approach reduce clinical complications and save the health care industry billions of dollars annually? The answer, quite simply, is maybe, but no one knows for sure. Although the concept may be attractive and this evolving area needs to be researched, good clinical data to support the use of genetic testing for warfarin dosing are not yet available. Also, it is important to realize that the impressive conclusions from the Brookings report mentioned above are based on supposition and projections, not on solid clinical outcome data. Furthermore, some experienced clinicians question whether genetic testing adds significantly to the information one may discern by carefully monitoring the international normalized ratio (INR) and by taking into consideration the numerous patient-specific factors that influence warfarin dosing requirements, such as age, underlying disease states, and concomitant drugs.

From the College of Pharmacy, University of Texas at Austin, Austin, Texas (Dr. Bussey); the Division of Pharmacotherapy, Clinical Pharmacy Programs, University of Texas Health Science Center, San Antonio, Texas (Dr. Bussey), ClotCare Online Resource (www.clotcare.com), San Antonio, Texas (Dr. Bussey and Ms. Walker); the School of Pharmacy, University of Washington, Seattle, Washington (Dr. Wittkowsky); and the School of Medicine, Boston University, Boston, Massachusetts (Dr. Hylek).

Address reprint requests to Henry I. Bussey, Pharm.D., FCCP, FAHA, Clinical Pharmacy Programs, Mail Code 6220, The University of Texas Health Sciences Center at San Antonio, 7703 Floyd Curl Drive, San Antonio, TX 78229-3900.

Genetic Testing for Warfarin Dosing

Scientific Basis

The genetic tests in question are useful in assessing an individual patient's sensitivity to warfarin, as well as that patient's rate of warfarin metabolism.

The test that can estimate a patient's sensitivity to warfarin is referred to as the vitamin K epoxide reductase (VKORC1) test; *VKORC1* is the gene that codes for the enzyme that is the site of action where warfarin exerts its effect. Genetic testing can indicate whether the patient may be more sensitive or less sensitive to warfarin than "average." Patients who are found to have the sensitive genotype (often referred to as AA genotype) typically require a lower dose than average. Those who are expected to be resistant to the effects of warfarin are referred to as GG

genotype and typically require larger doses of warfarin to reach a therapeutic INR. Those with the usual genotype are referred to as AG genotype.

The test that can estimate a patient's rate of warfarin metabolism is referred to as the cytochrome P450 (CYP) 2C9 test (or just "2C9"); CYP2C9 refers to the particular liver enzyme that is primarily responsible for metabolizing the most active component of warfarin. Some patients have a genetic variation in the CYP2C9 enzyme so that they metabolize warfarin more slowly than usual. Patients who metabolize warfarin more slowly will continue to accumulate warfarin in the blood over a longer period of time, and their INR will take longer to reach a stable level. Slow metabolizers typically require a lower dose of warfarin than those who metabolize warfarin faster.

Potential Benefits

The argument in favor of genetic testing is that a clinician can use this information to achieve the correct INR sooner, to maintain the INR in range, and thus to prevent complications. For example, if the patient has the AA genotype (warfarin sensitive), the clinician would logically start therapy with a lower warfarin dose and/or would expect to see an earlier and larger than usual increase in the INR. For the GG genotype (warfarin resistant), the clinician might start with a larger than usual warfarin dose. For the patient with a *CYP2C9* variant that indicates slow metabolism, the clinician might start the patient with a typical initial warfarin dose to reach the target INR range quickly but would then need to adjust the dose downward over time to keep the INR in the target range as the patient accumulates more and more warfarin in the blood.

Another potential benefit suggested by some is that the results of these genetic tests may be put into a mathematic equation to determine a patient's specific warfarin dose and time needed to reach a steady state. An equation has been derived to use such genetic testing in order to predict the warfarin dose and the time required to reach a steady-state INR.[7]

Limitations and Risks

The use of genetic testing for warfarin dosing may have no significant benefit in practice and may significantly increase health care costs. In addition, the misuse of genetic information may increase the risk of warfarin therapy.

May Lack Significant Benefit in Practice

Although genetic testing has theoretical benefits, no good clinical data exist to show that providing clinicians with this genetic information will make any difference in practice. In fact, studies examining this issue have found genetic testing to account for only 39–56% of the variability in the warfarin dose.[8–14] One thorough study from the University of Florida used a linear regression model that included the *CYP2C9* and *VKORC1* genetic testing together with various other patient-specific factors (e.g., weight, smoking status, factor X genotype, factor VII genotype, and vitamin K intake) and found that a combination of all the factors they used explained only 51.4% of the variability in warfarin dose.[14]

Careful monitoring of the INR (as is necessary for safe and effective therapy) along with other clinical observations (such as race) can provide the information necessary for optimal dosing. Before the *VKORC1* gene was identified, clinicians recognized that Asians (who tend to have the AA genotype) tend to be more sensitive to warfarin, and African-Americans (who tend to have the GG genotype) tend to be warfarin resistant. However, in the absence of a clear racial effect, if the patient has an early and rapid increase after starting a reasonable dose of warfarin, that patient is likely to have the sensitive (AA) *VKORC1* genotype. If the patient requires a prolonged period of time before the INR stabilizes, then that patient is likely to have a *CYP2C9* variant that leads to a slower rate of warfarin metabolism. In addition, careful monitoring of the INR allows the clinician to identify and assess the impact on the INR of several other important variables that are not identified with genetic testing. Such variables include (but are not limited to) smoking habits, use of alcohol, other medical conditions (such as heart failure), exercise routines, drug therapies, and routine vitamin K intake. Consequently, close monitoring of the INR may provide insight into the patient's probable genetic profile, as well as provide additional useful information of the composite effect of various other factors. Clearly, there is no substitute for close INR monitoring of warfarin-treated patients.

May Significantly Increase Health Care Costs

Another potential limitation of genetic testing for warfarin dosing is the cost associated with these tests. The genetic tests in question cost

approximately $250/test or $500 for both tests. The American Enterprise Institute–Brookings Joint Center report estimated that 2 million people in the United States begin warfarin therapy every year.[6] Consequently, the cost of performing genetic testing on these 2 million patients would be approximately $1 billion. Whether genetic testing adds a significant level of additional useful information and whether that additional level of useful information is worth the costs remain to be determined.

May Increase Risk of Warfarin Therapy through Misuse of Genetic Information

There is the potential for the misuse of genetic information to be dangerous to the patient's care. For example, if clinicians rely too heavily on equation-based estimates of warfarin dose, they may not follow their patients' INR as closely as they should.

In the experience of at least one author of this editorial (H.I.B.), the equation[7] can accurately predict the warfarin dose in some patients, but in other patients the required dose may be twice as large or only one half as large as the equation-calculated warfarin dose. If clinicians rely on the equation-calculated dose and time to steady state, they may conclude that the usual frequency of INR monitoring is not needed. In those patients in whom the estimated dose is only one half of what is actually needed—or twice the dose that is appropriate for a given patient—the result of less frequent INR monitoring could be a catastrophic blood clot or bleeding complication. For example, if the equation tells the clinician that the daily dose will be 4.63 mg and that the patient will be at steady state in 12–15 days, then the clinician may rely on that information and not monitor the INR as often as is necessary. In reality, that same patient may require only 2.32 mg/day and may reach a steady state in 10 days. If the patient is receiving 4.63 mg/day and the INR is not followed closely, the INR could be dangerously high after 8 days of dosing. It is critical for clinicians to understand that the genetic-based dosing equations are assessing only two of the numerous factors that may influence the patient's response to warfarin.

Conclusion

The bottom line is that genetic testing for warfarin dosing may hold promise, but its time has not yet arrived. Clearly, more research is needed in this area, and solid clinical data demonstrating a clear benefit of such testing should be required before such testing is recommended on a routine basis.

References

1. **Palca J**. Genome project begins paying dividends to patients. National Public Radio. June 1, 2005. Available from http://www.npr.org/templates/story/story.php?storyId=467594. Accessed July 27, 2007.
2. **Western Pennsylvania Hospital**. Genetic test offered at West Penn lessens dangers of Coumadin/warfarin. March 13, 2007. Available from http://www.wpahs.org/wph/news/index.cfm?hera=590A5150&zeus=5B05. Accessed July 27, 2007.
3. **Johnson L**. DNA tests to determine warfarin dose. Washington Post. The Associated Press. January 12, 2007. Available from http://www.washingtonpost.com/wp-dyn/content/article/2007/01/12/AR2007011201331.html. Accessed July 27, 2007.
4. **Winslow R, Matthews A**. New genetic tests boost impact of drugs. The Wall Street Journal. December 21, 2005. Available from http://www.post-gazette.com/pg/05355/ 625848.stm. Accessed July 27, 2007.
5. **U.S. Food and Drug Administration**. Critical path initiative: warfarin dosing. Available from http://www.fda.gov/oc/initiatives/criticalpath/warfarin.html. Accessed July 27, 2007.
6. **McWilliam A, Lutter R, Nardinelli C**. Health care savings from personalizing medicine using genetic testing: the case of warfarin. American Enterprise Institute–Brookings Joint Center. November 2006. Available from http://aei-brookings.org/publications/abstract.php?pid=1127. Accessed July 27, 2007.
7. **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;106:2329–33.
8. **Gage BF, Eby C, Milligan PE, Banet GA, Duncan JR, McLeod HL**. Use of pharmacogenetics and clinical factors to predict the maintenance dose of warfarin. Thromb Haemost 2004;91:87–94.
9. **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;3:137–45.
10. **Millican E, Jacobsen-Lenzini PA, Milligan PE, et al**. Genetic-based dosing in orthopedic patients beginning warfarin therapy. Blood 2007;110:1511–15.
11. **Voora D, Eby C, Linder MW, et al**. Prospective dosing of warfarin based on cytochrome P-450 2C9 genotype. Thromb Haemost 2005;93:700–5.
12. **Wadelius M, Chen LY, Eriksson N, et al**. Association of warfarin dose with genes involved in its action and metabolism. Hum Genet 2007;121:23–34.
13. **WarfarinDosing.org**. Home page. Available from http://www.warfarindosing.org. Accessed July 27, 2007.
14. **Aquilante CL, Langaee TY, Lopez LM, et al**. Influence of coagulation factor, vitamin K epoxide reductase complex subunit 1 and cytochrome P450 2C9 gene polymorphisms on warfarin dose requirements. Clin Pharmacol Ther 2006;79:291–302.

Academic performance of ethnic minorities in medical school

May be adversely affected by negative stereotyping

FOTOLIA

RESEARCH, p 611, p 615

Phyllis Carr associate dean for student affairs, Boston University School of Medicine, Boston, MA 02493, USA
plcarr@bu.edu
Jonathon Woodson associate dean for diversity, Boston University School of Medicine, Boston, MA 02493, USA

Competing interests: None declared.

Provenance and peer review: Commissioned; not externally peer reviewed.

Cite this as: *BMJ* 2008;337:a1094
doi: 10.1136/bmj.a1094

Increasing the number of doctors in the workforce who come from minority groups has been proposed as a way to tackle the health disparities of minority populations. Several themes have arisen regarding the education of such doctors, including low numbers of students from ethnic minorities applying for medical school, worse prior preparation in the sciences and humanities, and underachievement in their medical education.[1]

The qualitative study by Woolf and colleagues (doi: 10.1136/bmj.a1220) provides important insights into one aspect that affects clinical education in the United Kingdom—how ethnic stereotyping can add to a downward trend in performance.[2] This reflects findings on how race affects the evaluation of African-American students in the United States, including the nature of the student-teacher interaction and the biases that may affect the evaluation of students and the educational process.[1] Differences in the expectations and treatment of students that stem from pervasive negative stereotypes adversely affect their learning and self confidence. This seems to be particularly so during the clinical parts of a medical education, when differences in cultural formality, linguistic differences in accents, and communication styles can affect not only the interaction between teacher and student, but the ability of the student to connect with the patient and obtain a comprehensive medical history. When students from a minority group are perceived by faculty to be shy, quiet, reserved, not engaged in their education, and following parental motivation rather than an innate passion for medicine, their ability to excel is greatly reduced.

Woolf and colleagues suggest that getting to know the individual student on a personal level is an effective means to tackle "stereotype threat." For a profession such as medicine, we should not only teach students, but also mentor them. Studies in business have shown that people from minority groups who received mentoring on instructional development skills reached a plateau in middle management, whereas those whose mentors taught them broader developmental skills—which tackled negative stereotypes, public scrutiny, difficulty with role modelling, and peer resentment—achieved higher executive positions.[3]

Knowing students as individuals is important in any form of education, but particularly in medicine, where collegial interactions and those with patients are so key to the quality of care. This can also help in tackling the "hidden curriculum"—the extracurricular influences that a medical student is exposed to in medical schools—which often negates the professionalism and values that we need to instil in our students.[4] Similarly, we need to be aware of the "mental models" that we harbour—the picture that we spontaneously recall when approached by a person of a particular ethnicity—so that we can reconsider how we develop awareness of the attitudes and perceptions that influence our thoughts and interactions with people from minority groups.[5] We live in a world of increasing diversity, and we need to understand and appreciate differences in race, ethnicity, and culture.

Woolf and colleagues also report sex differences that reflect another dimension of our ability to provide excellence in medical education. In the medical school where their study took place, women were largely aggregated in general practice (five women), whereas men were more likely to be consultant physicians (10 men, two women), or consultant surgeons (five men, one woman). Why aren't the numbers of men and women in each specialty more equal? The disparity reflects the glass ceiling affecting women and their advancement in academic medicine, including specialties in which women are under-represented, such as surgery.

When questioned about the lack of engagement of students, female doctors questioned their teaching ability and attempted to find out the reason for the lack of engagement and tried to get to know their students. Male surgeons indicated that they decreased their teaching or made it more difficult for students who were less engaged. Unfortunately the authors did not comment further on the potential reasons behind these differences between the sexes. Ethnic and racial differences are important, but gender differences are equally pervasive.

The study provides important new insights and raises other questions about the medical curriculum. Should students be given the opportunity to develop and receive feedback on the quality of their social interactions? Should culturally determined patterns of communication be altered to optimise interaction with patients or teachers? Both of these questions can be dealt with by ensuring the cultural competence of our medical graduates through more formal curriculums and evaluation of these important skills. We also need

better ways to evaluate the motivations of candidates applying to medical schools.

1 Edelin KC; Ugholue A. Evaluation of an early medical school selection program for underrepresented minority students. *Acad Med* 2001;76:1056-9.
2 Woolf K, Cave J, Greenhalgh T, Dacre J. Ethnic stereotypes and the underachievement of UK medical students from ethnic minority groups: qualitative study. *BMJ* 2008;337:a1220.
3 Thomas DA. The truth about mentoring minorities. Race matters. *Harvard Bus Rev* 2001;79:98-107.
4 Haidet P, Stein HF. The role of the student-teacher relationship in the formation of physicians. The hidden curriculum as process. *J Gen Intern Med* 2006;21(suppl 1):S16-20.
5 Senge P, Kleiner A, Roberts C, Ross R, Roth G, Smith B. *The dance of change: the challenges to sustaining momentum in learning organizations.* New York: Doubleday, 1999.

The self controlled case series method

A way to study the relation between antipsychotic drugs and stroke

TEK IMAGE/SPL

RESEARCH, p 616

Heather Whitaker lecturer in statistics, Department of Mathematics and Statistics, Open University, Milton Keynes MK7 6AA
h.j.whitaker@open.ac.uk

Competing interests: None declared.
Provenance and peer review: Commissioned; not externally peer reviewed.

Cite this as: *BMJ* 2008;337:a1069
doi: 10.1136/bmj.a1069

The study by Douglas and Smeeth (doi: 10.1136/bmj.a1227) uses the self controlled case series method to study the association between exposure to antipsychotics and the risk of stroke.[1] The study found that use of any antipsychotic agent significantly increased the risk of stroke (relative risk 1.73, 95% confidence interval 1.60 to 1.87). The risk of stroke in people with dementia taking any antipsychotic was higher (3.50, 2.97 to 4.12) than in people without dementia taking similar medication (1.41, 1.29 to 1.55).

The self controlled case series method, or case series method for short, can be used to study the association between an acute event and a transient exposure using data only on cases; no separate controls are needed.[2]

The method uses exposure histories that are retrospectively ascertained in cases to estimate the relative incidence. That is, the incidences of events within risk periods—windows of time during or after experiencing the exposure when people are hypothesised to be at greater risk—relative to the incidences of events within control periods, which includes all time before the case experienced the exposure and after the risk has returned to the baseline value.

This method has two major advantages. Firstly, no controls are needed, which reduces the time needed to carry out a study, as well as the costs. Secondly, all fixed multiplicative confounders that do not vary over the study periods and that act proportionally on the baseline risk are controlled for implicitly. Another advantage is that in certain circumstances, when the risk periods are short compared with the total observation time, the case series method is almost as efficient as the cohort method with the same number of cases. Such a situation often arises in vaccine safety studies, for which this method has been used most often to date.

For many researchers, the main appeal of the self controlled case series method is the implicit control of fixed confounders. This was the case for Douglas and Smeeth, who stated that the underlying cardiovascular risk for people prescribed and not prescribed antipsychotic drugs differs in ways that are difficult to quantify and control for. Databases often do not include as much information on potential confounders as researchers would like, making the case series design attractive for studies using database data, as was the case in the present study.[3]

Whereas fixed confounders are controlled for, confounders that vary with time are not, although it is possible to allow for them explicitly. Age effects are almost always included in a case series analysis; the increasing incidence of stroke with age was allowed for by including an age effect in five year age bands. Results can sometimes be sensitive to the choice of age groups, and a semi-parametric version of the method—which avoids the need to specify the age groupings at the cost of increased computational time— is available.[4] For large data sets, a more practical approach is to undertake sensitivity analyses with different age groupings.

The largest limitation of the self controlled case series method is that the probability of exposure must not be altered by a previous event. If stroke were a contraindication to antipsychotic drugs the relative incidence would be biased upwards. Is it possible that the occurrence of a stroke would lead to people being more or less likely to be prescribed antipsychotic drugs? This study looked at patients only in the time period before the end of 2002 because this was when concerns about the possible effects of antipsychotic drugs first emerged, hence minimising the chance that any such bias would be present. Events that can result in death, as is the case with stroke, can also introduce bias—patients must be alive to receive a prescription. It is unclear what effect this would have in the present study, although the bias was shown to be small in a study of myocardial infarction.[4]

Several studies have carried out both a case-control analysis and a case series analysis.[5] The sources of bias are different in the two types of studies, and comparing results from each might provide insight into how these biases affect the results or might increase confidence in study conclusions.

Case series studies are relatively straightforward to perform. Observation periods are defined for each case by fixing age and time boundaries for the study, exposure histories are ascertained, and age groups and risk periods are defined. Perhaps the most difficult part is choosing how to define the risk periods. These should be defined a priori, on the basis of the study hypotheses, previous studies, and biological mechanisms. Where uncertainty exists, several contiguous periods can be used. For example, Douglas and

Original Report: Patient-Oriented, Translational Research

American Journal of Nephrology

Am J Nephrol 2008;28:354–360
DOI: 10.1159/000111829

Received: June 14, 2007
Accepted: October 15, 2007
Published online: November 29, 2007

Chronic Kidney Disease Is Associated with Angiographic Coronary Artery Disease

Michel Chonchol[a, b] Jeff Whittle[c] Angela Desbien[b] Michelle B. Orner[d]
Laura A. Petersen[e] Nancy R. Kressin[d, f]

[a]Division of Renal Diseases and Hypertension and [b]Department of Medicine, University of Colorado Health Sciences Center, Denver, Colo., [c]Primary Care Division, Clement J. Zablocki VA Medical Center, Milwaukee, Wisc., [d]Center for Health Quality, Outcomes and Economic Research, Bedford VA Medical Center, Bedford, Mass., [e]Division of Health Policy and Quality, Houston Center for Quality of Care and Utilization Studies, a Health Services Research and Development Center of Excellence, Houston VA Medical Center, and Section of Health Services Research, Baylor College of Medicine, Houston, Tex., and [f]Section of General Internal Medicine, Boston University School of Medicine, Boston, Mass., USA

Key Words
Chronic kidney disease · Coronary artery disease · Coronary angiography

Abstract
Background/Aims: Patients with chronic kidney disease (CKD) have a dramatically increased risk for cardiovascular mortality. Few prior studies have examined the independent association of CKD with coronary anatomy. ***Methods:*** We evaluated the relationship between CKD and severe coronary artery disease (CAD) in 261 male veterans with nuclear perfusion imaging tests suggesting coronary ischemia. We used chart review and patient and provider interviews to collect demographics, clinical characteristics, and coronary anatomy results. We defined CKD as an estimated glomerular filtration rate (eGFR) <60 ml/min/1.73 m^2, based on the creatinine obtained prior to angiography. We defined significant coronary obstruction as at least one 70% or greater stenosis. We used logistic regression to determine whether CKD was independently associated with significant coronary obstruction. ***Results:*** The likelihood of CAD increased monotonically with decreasing eGFR, from 51% among patients with eGFR ≥90 ml/min/1.73 m^2 to 84% in those with eGFR <30 ml/min/1.73 m^2 ($p = 0.0046$). Patients with CKD were more likely than those without CKD to have at least one significant coronary obstruction (75.9 vs. 60.7%, $p = 0.016$). Patients with CKD also had more significant CAD, that is, were more likely to have three-vessel and/or left main disease than those without CKD (34.9 vs. 16.9%, $p = 0.0035$). In logistic regression analysis, controlling for demographics and comorbidity, CKD continued to be independently associated with the presence of significant CAD ($p = 0.0071$). ***Conclusion:*** CKD patients have a high prevalence of obstructive coronary disease, which may contribute to their high cardiovascular mortality.

Introduction

Patients with chronic kidney disease (CKD) are at high risk for cardiovascular morbidity and mortality [1, 2]. In fact, CKD patients are far more likely to die from cardiovascular disease (CVD) than progress to end-stage renal disease [3]. Recent estimates suggest that more than 20 million people have CKD in the USA alone [4]. Furthermore, the CKD and dialysis populations are growing rapidly and are expected to exceed 30 million and 650,000 respectively in the USA by 2010 [5]. While CVD risk fac-

KARGER
Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

0250–8095/08/0282–0354$24.50/0

Accessible online at:
www.karger.com/ajn

Michel Chonchol, MD
University of Colorado Health Sciences Center
Division of Renal Diseases and Hypertension, Box C-281
Denver, CO 80262 (USA)
Tel. +1 303 399 6997, Fax +1 303 399 3131, E-Mail Michel.Chonchol@uchsc.edu

tors in CKD have not been well defined, the increased cardiovascular risk may in part be due to both traditional and non-traditional risk factors associated with kidney dysfunction, including hypertension, diabetes and inflammation [6, 7].

Autopsy [8] and clinical [9] studies have documented a higher prevalence of coronary artery plaques in patients with end-stage renal disease. Therefore, the increased risk of CVD in patients requiring chronic dialysis is considered to be related to advanced atherosclerosis [10]. Although the increased risk of cardiovascular events among persons with CKD not requiring dialysis [2] is well established, it is not clear what therapeutic steps should be taken, since the mechanism for the increase in risk is not clear. If there were an increase in rates of angiographic coronary artery disease (CAD) among persons with CKD, a more aggressive approach to the assessment of coronary anatomy might be indicated, despite the risks to kidney function associated with the use of angiographic contrast material. However, existing literature is inconclusive regarding whether the increase in risk of events seen in persons with CKD not requiring renal replacement therapy is associated with an increase in coronary obstructive disease. One meta-analysis suggested that there was no increase in rates of obstructive disease in the presence of CKD [11], while smaller studies in selected populations have suggested there is a relationship [12, 13].

A disconnect between coronary anatomy and CVD events is plausible, since epicardial vessel stenosis is a far from perfect predictor of adverse outcomes [14, 15]. We therefore conducted the present study in a well-characterized group of men with and without CKD undergoing coronary angiography for suspected coronary disease following a positive myocardial perfusion imaging study. We sought to determine whether CKD was associated with increased coronary obstructive disease among participants in the Cardiac Decision Making Study (CDMS) [16] after controlling for a comprehensive set of clinical variables including an estimate of the likelihood of coronary obstruction made by the managing physician most involved in the decision to proceed to coronary angiography.

Methods

Study Population

The CDMS [16] is an observational cohort study of self-identified white or African-American veterans who underwent non-invasive testing for coronary obstructive disease at one of five participating VA medical centers between August 1999 and January 2001. Study personnel screened all veterans who had a nuclear imaging study that suggested coronary ischemia, according to the local interpreter. These studies were performed as part of routine clinical care. Thus, they included a variety of techniques (resting and delayed, various exercise protocols, and various pharmacological stressors) and radionuclides (e.g., single and dual isotope) in keeping with local practice patterns. Details of the recruitment process have been published [17, 18], but we present a brief summary.

We screened 5,278 patients who had a nuclear imaging study, of whom 2,335 (44%) had a positive study. Of these, we excluded 981 for the following reasons: 456 patients (19%) could not be contacted to enroll in the study; 209 (9%) had had a cardiac procedure in the preceding 6 months; 102 (4%) were not African-American or white; 78 (3%) had impaired mental status; 32 (1%) were in another research study determining their cardiac treatment, and 104 (4%) were excluded for miscellaneous other reasons (e.g., the nuclear imaging study was conducted for a compensation and pension evaluation, the patient's hearing was impaired, or the patient was not a veteran) [16]. Of the remaining 1,354 patients with positive imaging studies, 329 overtly refused, failed to return their informed consent, or failed to return mailed questionnaires. Thus, 1,025 (75.7%) persons out of 1,354 eligible veterans who were contacted agreed to participate. Of these, 318 subjects had coronary angiography within 90 days following their imaging study. For the present analysis we included only male veterans who had a serum creatinine within 180 days preceding the date of the angiogram (n = 261). We excluded 57 subjects due to (1) no serum creatinine recorded (n = 32), (2) serum creatinine recorded on a date after the angiogram (n = 18), and (3) female gender (n = 7). We excluded women because there were too few to allow for stratified analysis.

To examine the representativeness of the sample, we compared the 261 patients with positive nuclear imaging study results who were included in the cohort to those male patients who agreed to participate but were not included as they did not have a coronary angiography within 90 days of their imaging study (n = 747). Since African-Americans were less likely to undergo coronary angiography [16], there were fewer African-Americans in the cohort (18% African-Americans in the cohort vs. 24% in the excluded patients, p = 0.0309). There were no differences in age and prevalence of hypertension, diabetes, history of prior revascularization, congestive heart failure (CHF) and CKD between these groups. In addition, the estimated glomerular filtration rate (eGFR) was similar between both groups (68.9 ± 23.4 ml/min/1.73 m^2 in the cohort vs. 69.3 ± 24.8 ml/min/1.73 m^2, p = NS). Furthermore, the 50 individuals excluded because the serum creatinine was not available or recorded after the coronary angiogram were less likely to have diabetes (16 vs. 32.2%, p = 0.0216) and more likely to have a history of revascularization (50 vs. 30.5%, p = 0.0076) than those included in this analysis. The prevalence of all other cardiovascular risk factors was similar between these two groups.

Data

We reviewed the medical records of each study respondent, obtaining records for non-VA care where possible. Trained nurses who were blinded to study aims abstracted patient demographics, cardiac symptoms, and past medical history (including prior myocardial infarction or prior coronary revascularization, angina, CHF, diabetes, hypertension, kidney dysfunction or chronic

Table 1. Baseline characteristics of the cohort by eGFR

eGFR, ml/min/1.73 m^2	81.3 ± 15.1	42.2 ± 13.7	<0.0001
Age			0.0003
20–54 years	55 (30.9)	9 (10.8)	
55–64 years	54 (30.3)	26 (31.4)	
≥65 years	69 (38.8)	48 (57.8)	
Site of care			0.1485
Atlanta	30 (16.8)	17 (20.5)	
Durham	35 (19.7)	25 (30.1)	
Houston	36 (20.2)	18 (21.7)	
Pittsburgh	45 (25.3)	14 (16.9)	
St. Louis	32 (18.0)	9 (10.8)	
White race	148 (83.2)	67 (80.7)	0.6323
Prior myocardial infarction	55 (31.4)	36 (44.4)	0.0430
Prior revascularization	51 (29.0)	28 (33.7)	0.4377
Maximal medical therapy	67 (37.6)	40 (48.2)	0.1065
Angina	140 (79.10)	51 (62.20)	0.0040
Hypertension	135 (76.3)	76 (91.6)	0.0033
Diabetes	53 (29.8)	31 (37.4)	0.2226
Congestive heart failure	18 (10.2)	37 (44.6)	<0.0001
Chronic obstructive pulmonary disease	39 (22.0)	22 (26.5)	0.4276
Nuclear study			0.9070
Nuclear study, high risk	90 (51.7)	43 (53.8)	
Nuclear study, moderate risk	74 (42.5)	30 (37.5)	
Nuclear study, low risk	10 (5.8)	7 (8.8)	
Physician estimated likelihood of obstructive CAD, %	82.4 ± 17.7	85.8 ± 14.2	0. 1291

Values expressed as number of participants (%). Values presented with ± are means ± SD.
CAD = Coronary artery disease; eGFR = estimated glomerular filtration rate. p values are for the difference between groups.

obstructive lung disease). The abstractors used the official coronary angiogram report to determine the presence of obstruction in each of the major coronary systems, and whether there was obstruction of the proximal left anterior descending artery. As an indication of the extent to which medical therapy had been maximized for each patient, we used the American College of Cardiology/American Heart Association guidelines for coronary angiography and the management of patients with chronic stable angina [19, 20]. Thus, we defined maximal medical therapy as antiplatelet therapy, sublingual nitroglycerin and at least one of the following: β-blockers, calcium channel blockers, or long-acting nitrates.

As the managing physicians had information on the history, cardiovascular risk factors and the results of the nuclear medicine study, they were asked to estimate the probability that the patients had clinically significant CAD. To assess this perception, we asked, 'On a scale from 0–100%, please estimate the probability of coronary artery disease (CAD) in this patient (70% or more narrowing of an epicardial artery)'. This question was asked prior to patients undergoing a coronary angiogram. We obtained the managing physician's estimate of the likelihood of significant CAD for 216 (82.8%) of the participants.

We classified coronary obstruction as severe if either the left main coronary artery or all three major coronary systems had a stenosis of 70% or greater. We classified non-severe obstructions as moderate if the proximal left anterior descending artery was involved (stenosis of 70% or greater), and mild if it was not, but if there was at least one coronary obstruction of greater than 70%. We classified coronary obstruction as none if there was no obstruction of greater than 70% [21].

Two physicians, a board-certified general internist (J.W.) and a cardiology fellow, classified the severity of each nuclear imaging study based on review of the official report, if available (n = 254, 97%). We categorized the risk of severe coronary obstruction as low, moderate or high, using a modification of the methods of Bateman et al. [22]. In this method, patients with reversible lesions in the distribution of left anterior descending coronary artery or in both the right coronary artery and left circumflex artery were considered to be at high risk, as were patients with increased lung uptake or transient ischemic dilatation with exercise or pharmacologic stress. Patients with reversible lesions in just one of the right coronary artery or left circumflex artery were considered to be at moderate risk. Patients whose defects were very small or minimally reversible were considered to be at low risk. Disagreements were resolved by consensus.

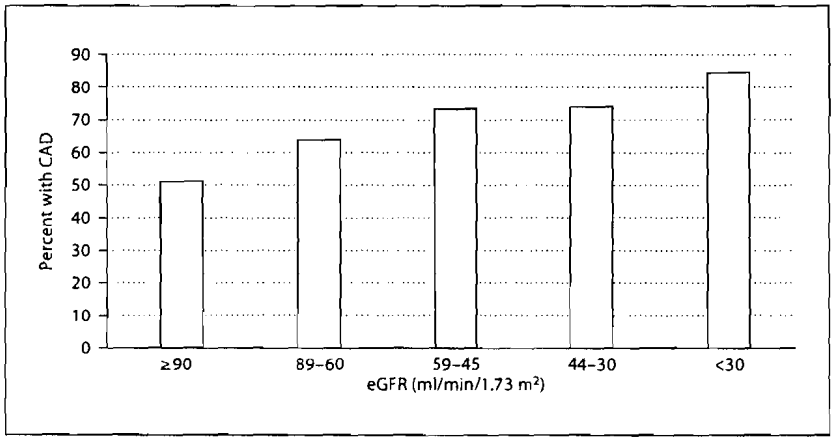


Fig. 1. Relationship of prevalence of CAD to estimated glomerular filtration rate. CAD was defined as at least one coronary artery with 70% or greater stenosis on coronary angiography.

Definition of Chronic Kidney Disease

Recent published guidelines recommend estimating kidney function using a formula derived by the Modification of Diet in Renal Disease (MDRD) study group [23]. This equation expresses GFR in ml/min per 1.73 m^2 body surface area (BSA) and is calculated as: 186 × (serum creatinine$^{-1.154}$) × (age$^{-0.203}$) × 1.21 (if black) × 0.742 (if female) where serum creatinine is measured in mg/dl and age is in years. We used the last serum creatinine in the database that was performed prior to the coronary angiogram but within 180 days of the procedure. We defined CKD as an eGFR <60 ml/min/1.73 m^2. Serum creatinine was assessed by spectrophotometry analysis using the modified kinetic Jaffé reaction [24] at all sites.

Statistical Analysis

We examined the bivariate association of CKD with coronary obstruction using a χ^2 test for trend. We then dichotomized coronary obstruction as no obstructive disease versus any for further analysis. We used χ^2 or t tests as appropriate to compare demographics and clinical characteristics between patients with and without obstructive disease. We examined age as both a continuous and a categorical variable, but present our results using the categorical treatment, as results were similar.

We used logistic regression to analyze the association between risk factors and presence of coronary obstruction controlling for confounding factors. In the initial multivariable model we included physician estimate of the likelihood of significant CAD, prior revascularization, and nuclear study classification adjusted for site. This model was chosen because it was known to be a strong predictor of CAD in the overall CDMS cohort [16]. We then added CKD, followed by age, race, diabetes, hypertension and CHF. The final variables were chosen on the basis of biological plausibility for the relationship of CKD to CAD. Site of care was treated as a random effect. This approach accounts for the possibility that patients within the same site might share similar, unmeasured, characteristics that are associated with the presence of coronary obstruction [25]. All analyses were conducted using SAS statistical software (SAS Inc., Cary, N.C., USA). This study was approved by the human studies subcommittee of the VA medical centers where data collection took place, and by the study coordinating center.

Results

Of the 261 participants included in the analytic cohort, a total of 83 (31.8%) subjects had CKD (eGFR <60 ml/min/1.73 m^2) at baseline. Only 3 (1.1%) patients had an eGFR <15 ml/min/1.73 m^2. The characteristics of participants with and without CKD are shown in table 1. Compared with those who did not have CKD, participants with CKD were older and more likely to have a history of myocardial infarction, hypertension or CHF. Patients with CKD were less likely to have angina reported in the medical record. Physician estimates of the likelihood of CAD were similar in patients with and without CKD, as was the severity of ischemia on the index nuclear imaging study.

A significant trend for the presence of CAD was noticed when using decreasing levels of eGFR as an ordinal variable (p = 0.0046) (fig. 1). Patients with CKD were more likely to have at least one significant (≥70% stenosis) coronary obstruction than those without CKD (75.9 vs. 60.7%, p = 0.016). In addition, CKD patients had more extensive CAD than non-CKD patients (p = 0.0035) (fig. 2). Of note, CKD patients were more likely to have severe obstructive CAD (34.9 vs. 16.9%) and less likely to have no significant obstructions (39.3 vs. 24.1%).

Table 2 shows the results of our multivariable analysis for the presence of coronary obstruction, which adjusts for clustering of observations within each site of care. After adjusting for site, demographics, and clinical factors we found that CKD remained significantly associated with the presence of CAD. Other clinical variables that were associated with the presence of significant CAD included prior revascularization, nuclear imaging results

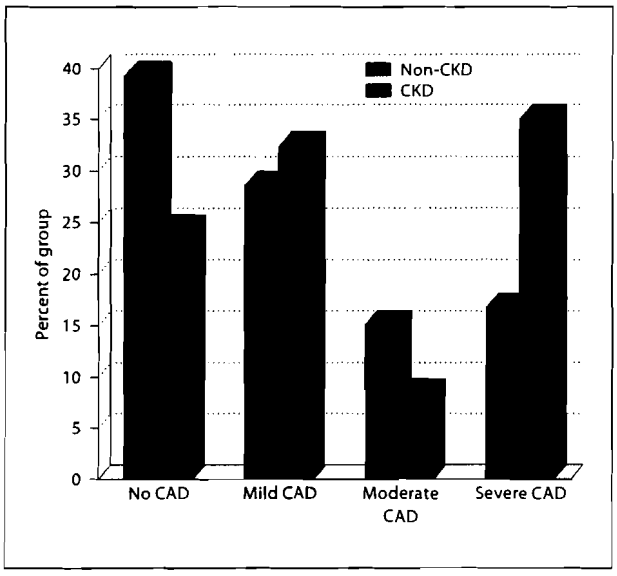


Fig. 2. Relationship of kidney function to severity of CAD. The distribution of severe CAD defined as a significant obstruction (≥70%) of left main or all three coronary artery systems is significantly different between patients with and without CKD, p = 0.0035. CKD was defined as an eGFR <60 ml/min/1.73 m^2.

Table 2. Multiple logistic regression analysis for having at least one 70% or greater stenosis on coronary angiography, adjusted for site of care

Age group 20–54 vs. ≥65	1.03 (0.43; 2.46)	0.9556
Age group 55–64 vs. ≥65	1.11 (0.52; 2.37)	0.7918
White race	1.60 (0.69; 3.72)	0.2747
Diabetes	0.87 (0.42; 1.79)	0.6973
Hypertension	0.81 (0.36; 1.86)	0.6240
Congestive heart failure	0.45 (0.18; 1.12)	0.0863
Prior revascularization	3.08 (1.33; 7.14)	0.0092
Nuclear imaging study results		
low vs. high risk	0.20 (0.04; 0.98)	0.0473
moderate vs. high risk	0.47 (0.23; 0.96)	0.0370
Physician estimated risk	1.03 (1.01; 1.05)	0.0116
Chronic kidney disease	3.16 (1.37; 7.27)	0.0071

and physician estimated risk of CAD. Although age, race, history of diabetes mellitus, hypertension and CHF were not significantly related to the presence of CAD in this analysis, we left them in the model because of their known association with both CKD and CAD. When age was used as a continuous variable similar results were obtained.

Discussion

In this study of male veteran patients with nuclear imaging studies suggesting ischemia, we report a significant relationship between CKD and presence of significant CAD. In particular, patients with an eGFR <60 ml/min/1.73 m^2 were more likely to have significant coronary obstruction than patients with higher eGFR after controlling for demographic and clinical variables. We also found the risk of significant coronary obstruction increased progressively with decreasing kidney function. In addition to documenting the relationship of coronary anatomy to kidney function, these results provide further support for the growing literature showing a strong association between cardiovascular mortality and CKD [2].

Other important predictors of the presence of severe CAD included: prior revascularization, results of nuclear imaging study and the managing physician's estimate of the risk of CAD. The managing physicians had a tendency to overestimate the likelihood of coronary disease in patients with (estimate 85.8%, actual 76%) and without (estimate 82.4%, actual 60.7%) kidney disease, although the estimate was more comparable to the actual percentage of angiographic CAD in the group with kidney disease. The managing physician assessment of the likelihood of obstructive CAD was based mainly on the available clinical data on chart review, which included information regarding kidney function. These findings suggest that the presence or absence of kidney disease did not appear to influence the managing physician estimated risk of CAD in our study and imply that physician awareness of the association between CKD and angiographic CAD needs to be enhanced.

Relatively few studies have examined the association between kidney function and coronary anatomy in patients with moderate to severe kidney dysfunction who were not yet on dialysis [11–13]. These findings are contradictory. In a recent meta-analysis of intravascular ultrasound studies, Nicholls et al. [11] compared 176 CKD patients not yet on dialysis (mean eGFR 53 ± 6 ml/min/1.73 m^2) with 813 patients with an eGFR >60 ml/min/1.73 m^2 (mean eGFR 77 ± 13 ml/min/1.73 m^2) with known CAD. These investigators found that the total atheroma volume in subjects with CKD did not differ from subjects with an eGFR >60 ml/min/1.73 m^2. In addition, no differences in progression rates of atherosclerotic dis-

ease were observed between these two groups of patients, suggesting that the increased incidence of cardiovascular events reported in patients with CKD may result from factors other than atherosclerotic burden [11].

Our findings are similar to those from a cross-sectional study of 56 women with symptomatic CVD and mild renal insufficiency who were referred for coronary angiography in the Women's Ischemia Syndrome Evaluation (WISE) study [12]. In this study, women with mild kidney dysfunction had higher rates of significant angiographic CAD (≥50% diameter stenosis in ≥1 coronary artery) compared with women with normal kidney function (61 vs. 37%; $p < 0.001$). This association persisted independent of age, total and low density lipoprotein-cholesterol, systolic and diastolic blood pressures, diabetes, hypertension, smoking, dyslipidemia, and menopausal status. Their study cohort was younger, included only women and the kidney dysfunction was defined on the basis of abnormal serum creatinine (1.2–1.9 mg/dl). Our findings are also similar to those from a cross-sectional study of 30 asymptomatic stage 5 CKD patients of which 16 (53.3%) had significant coronary obstructive disease [13]. In contrast to our study, this study did not include patients with stage 3 and 4 CKD, the sample size was small and the cohort only included Asian patients.

The pathophysiologic mechanisms for the association between CKD and angiographic CAD are still unclear although multiple explanations have been proposed [26–28]. Traditional risk factors that have correlated with CVD in CKD patients include older age, male gender, smoking, diabetes, hypertension, dyslipidemia and left ventricular hypertrophy resulting in abnormal left ventricular wall motion. Other more non-traditional cardiovascular risk factors associated with uremia secondary to progressive kidney disease include elevated serum calcium-phosphorus product, secondary hyperparathyroidism, hypoalbuminemia, hyperhomocysteinemia, increased circulating levels of proinflammatory cytokines, and oxidative stress which have all been associated with abnormal vascular biology in patients with CKD [26, 27].

Despite the comprehensive nature of the dataset, there are several limitations to consider in this analysis. First, our definition of kidney disease was based solely on information available from review of the medical record. Thus, we defined CKD based on an eGFR derived from a single serum creatinine determination within 180 days prior to the coronary angiogram, rather than on a serum creatinine on the day of the procedure or a direct measurement of kidney function like iothalamate clearance. Additionally, the creatinine value we did use could have been influenced by cardioprotective medications or clinical status. Furthermore, we did not collect data regarding the duration or cause of CKD or other signs of kidney disease such as microalbuminuria or overt proteinuria. Thus, it is likely that many of the subjects we classified as not having CKD, did indeed have some form of kidney disease, and that some of the participants we classified as having CKD had only had it for a relatively brief period of time. This type of misclassification, if random, would tend to cause us to underestimate the magnitude of the relationship between CKD and obstructive CAD. Thus, the fact that we observed a significant association suggests that the relationship is a strong one.

Second, the fact that we studied an older, male, chronically ill population makes it difficult to extrapolate our results to other populations. However, the burden of CKD is high among the population studied, making it an important result nonetheless. Additional studies are needed to extend this finding to other populations. Third, since our study population was small, our estimate of the effect of CKD has wide confidence intervals. Fourth, because of the observational nature of the study, we only have data on the coronary anatomy of participants who underwent coronary angiography in the course of their clinical care. We do not know if the relationship of CKD to obstructive CAD would have been different among those persons who did not undergo coronary angiography.

Finally, data on other important traditional (e.g., lipid profile, smoking history) and non-traditional (e.g., C-reactive protein, homocysteine) cardiovascular risk factors were not available for the purpose of this analysis. However, we emphasize that we controlled for the managing physician's estimate of the likelihood of coronary obstruction. That physician would have been considering all the relevant, available, clinical information as they made the decision to refer the person to coronary angiography. Thus, our results suggest that CKD status does add information about the likelihood of obstructive CAD to the information currently used in clinical practice.

In summary, we have shown that angiographic CAD is associated with CKD in a male population consisting of patients with positive nuclear imaging studies. This association was independent of important traditional cardiovascular risk factors. These data emphasize the importance of increased efforts for the aggressive diagnoses and treatment of CAD in the CKD population. Future research should clarify the pathophysiological mechanisms of these findings.

Acknowledgements

The research reported here was supported by the Department of Veterans Affairs, Veterans Health Administration, Health Services Research and Development Service (ECV 97-022.2, N. Kressin, P.I.), and the American Heart Association and the Pharmaceutical Roundtable (9970113N, N. Kressin, P.I.). Dr. Kressin is a Research Career Scientist, Department of Veterans Affairs, Health Services Research & Development Service at the Edith Nourse Rogers Memorial Veterans Hospital. Dr. Petersen is a Robert Wood Johnson Foundation Generalist Physician Faculty Scholar (grant No. 045444), and an American Heart Association Established Investigator Awardee (grant No. 0540043N). 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

1 Menon V, Sarnak MJ: The epidemiology of chronic kidney disease stages 1 to 4 and cardiovascular disease: a high-risk combination. Am J Kidney Dis 1990;45:223–232.
2 Go AS, Chertow GM, Fan D, McCulloch CE, Hsu CY: Chronic kidney disease and the risks of death, cardiovascular events, and hospitalization. N Engl J Med 2004;351:1296–1305.
3 Keith DS, Nichols GA, Gullion CM, Brown JB, Smith DH: Longitudinal follow-up and outcomes among a population with chronic kidney disease in a large managed care organization. Arch Intern Med 2004;164:659–663.
4 Coresh J, Wei GL, McQuillan G, Brancati FL, Levey AS, Jones C, Klag MJ: Prevalence of high blood pressure and elevated serum creatinine level in the United States: findings from the third National Health and Nutrition Examination Survey (1988–1994). Arch Intern Med 2001;161:1207–1216.
5 Xue JL, Ma JZ, Louis TA, Collins AJ: Forecast of the number of patients with end-stage renal disease in the United States to the year 2010. J Am Soc Nephrol 2001;12:2753–2758.
6 Sarnak MJ, Levey AS: Cardiovascular disease and chronic renal disease: a new paradigm. Am J Kidney Dis 2000;35:S117–S131.
7 Muntner P, Hamm LL, Kusek JW, Chen J, Whelton PK, He J: The prevalence of nontraditional risk factors for coronary heart disease in patients with chronic kidney disease. Ann Intern Med 2004;140:9–17.
8 Ansari A, Kaupke CJ, Vaziri ND, Miller R, Barbari A: Cardiac pathology in patients with end-stage renal disease maintained on hemodialysis. Int J Artif Organs 1993;16:31–36.
9 Rostand SG, Kirk K, Rutsky EA: Dialysis-associated ischemic heart disease: Insights from coronary angiography. Kidney Int 1984;25:653–659.
10 Ma KW, Greene EL, Raji L: Cardiovascular risk factors in chronic renal failure and hemodialysis population. Am J Kidney Dis 1992;19:505–513.
11 Nicholls SJ, Tuzcu EM, Hsu A, Wolski K, Sipahi I, Schoenhagen P, Crowe T, Kapadia SR, Hazen SL, Nissen SE: Comparison of coronary atherosclerotic volume in patients with glomerular filtration rates < or = 60 versus >60 ml/min/1.73 m(2): a meta-analysis of intravascular ultrasound studies. Am J Cardiol 2007;99:813–816.
12 Reis SE, Olson MB, Fried L, Reeser V, Mankad S, Pepine CJ, Kerensky R, Merz CN, Sharaf BL, Sopko G, Rogers WJ, Holubkov R: Mild renal insufficiency is associated with angiographic coronary artery disease in women. Circulation 2002;105:2826–2829.
13 Ohtake T, Kobayashi S, Moriya H, Negishi K, Okamoto K, Maesato K, Saito S: High prevalence of occult coronary artery stenosis in patients with chronic kidney disease at the initiation of renal replacement therapy: an angiographic examination. J Am Soc Nephrol 2005;16:1141–1148.
14 Iskandrian AE, Verani MS: Nuclear imaging techniques; in Topol EJ (ed): Textbook of Cardiovascular Medicine. Philadelphia, Lippincott-Raven, 1998, pp 1367–1394.
15 Topol EJ, Nissen SE: Our preoccupation with coronary luminology. The dissociation between clinical and angiographic findings in ischemic heart disease. Circulation 1995;92:2333–2342.
16 Whittle J, Kressin NR, Peterson ED, Orner MB, Glickman M, Mazzella M, Petersen LA: Racial differences in prevalence of coronary obstructions among men with positive nuclear imaging studies. J Am Coll Cardiol 2006;47:2034–2041.
17 Kressin NR, Chang BH, Whittle J, Peterson ED, Clark JA, Rosen AK, Orner M, Collins TC, Alley LG, Petersen LA: Racial differences in cardiac catheterization as a function of patients' beliefs. Am J Public Health 2004;94:2091–2097.
18 Kressin NR, Clark JA, Whittle J, East M, Peterson ED, Chang BH, Rosen AK, Ren XS, Alley LG, Kroupa L, Collins TC, Petersen LA: Racial differences in health-related beliefs, attitudes, and experiences of VA cardiac patients: scale development and application. Med Care 2002;40(1 suppl):I72–I85.
19 Gibbons RJ, Chatterjee K, Daley J, Douglas JS, Fihn SD, Gardin JM, Grunwald MA, Levy D, Lytle BW, O'Rourke RA, Schafer WP, Williams SV, Ritchie JL, Cheitlin MD, Eagle KA, Gardner TJ, Garson A Jr, Russell RO, Ryan TJ, Smith SC Jr: ACC/AHA/ACP-ASIM guidelines for the management of patients with chronic stable angina a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee on Management of Patients With Chronic Stable Angina). J Am Coll Cardiol 1999;33:2092–2197.
20 Scanlon PJ, Faxon DP, Audet AM, Carabello B, Dehmer GJ, Eagle KA, Legako RD, Leon DF, Murray JA, Nissen SE, Pepine CJ, Watson RM, Ritchie JL, Gibbons RJ, Cheitlin MD, Gardner TJ, Garson A Jr, Russell RO Jr, Ryan TJ, Smith SC Jr: ACC/AHA guidelines for coronary angiography. A report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee on Coronary Angiography). Developed in collaboration with the Society for Cardiac Angiography and Interventions. J Am Coll Cardiol 1999;33:1756–1824.
21 Whittle J, Conigliaro J, Good CB, Hanusa BH, Macpherson DS: Black-white differences in severity of coronary artery disease among individuals with acute coronary syndromes. J Gen Intern Med 2002;17:867–873.
22 Bateman TM, O'Keefe JH Jr, Dong VM, Barnhart C, Ligon RW: Coronary angiographic rates after stress single-photon emission computed tomographic scintigraphy. J Nucl Cardiol 1995;2:217–223.
23 National Kidney Foundation: K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and Stratification. Am J Kidney Dis 2002;39:S1–S266.
24 Walser M: Assessing renal function from creatinine measurements in adults with chronic renal failure. Am J Kidney Dis 1998;32:23–31.
25 Localio AR, Berlin JA, Ten Have TR, Kimmel SE: Adjustments for center in multicenter studies: an overview. Ann Intern Med 2001;135:112–123.
26 McCullough PA: Why is chronic kidney disease the 'spoiler' for cardiovascular outcomes? J Am Coll Cardiol 2003;41:725–728.
27 Schrier RW: Role of diminished renal function in cardiovascular mortality: marker or pathogenetic factor? J Am Coll Cardiol 2006;47:1–8.
28 Shlipak MG, Fried LF, Crump C, Bleyer AJ, Manolio TA, Tracy RP, Furberg CD, Psaty BM: Elevations of inflammatory and procoagulant biomarkers in elderly persons with renal insufficiency. Circulation 2003;107:87–92.

Thromboembolic Consequences of Subtherapeutic Anticoagulation in Patients Stabilized on Warfarin Therapy: The Low INR Study

Nathan P. Clark, Pharm.D., Daniel M. Witt, Pharm.D., Thomas Delate, Ph.D., Melissa Trapp, Pharm.D., David Garcia, M.D., Walter Ageno, M.D., Elaine M. Hylek, M.D., and Mark A. Crowther, M.D., for the Warfarin-Associated Research Projects and Other Endeavors (WARPED) Consortium

Study Objective. To quantify the absolute risk of thromboembolism associated with a significant subtherapeutic international normalized ratio (INR) in patients with previously stable anticoagulation while receiving warfarin.

Design. Retrospective, matched cohort analysis.

Setting. Centralized anticoagulation service in an integrated health care delivery system.

Patients. A total of 2597 adult patients receiving warfarin from January 1998–December 2005; 1080 patients were in the low INR cohort and were matched to 1517 patients in the therapeutic INR cohort based on index INR date, indication for warfarin, and age.

Measurements and Main Results. Stable, therapeutic anticoagulation was defined as two INR values, measured at least 2 weeks apart, within or above the therapeutic range. The low INR cohort included patients with a third INR value of 0.5 or more units below their therapeutic range. The therapeutic INR cohort included patients with a third therapeutic INR value and no INR value 0.2 or more units below their target INR range in the ensuing 90 days. The primary outcome was anticoagulation-related thromboembolism during the 90 days after the index INR. Secondary outcomes were times to the first occurrence of anticoagulation-related complications (bleeding, thromboembolism, or death) in the 90 days after the index INR. Four thromboembolic events (0.4%) occurred in the low INR cohort and one event (0.1%) in the therapeutic INR cohort ($p=0.214$). The differences in the proportions of thromboembolism, bleeding, or death were not significant between the cohorts ($p>0.05$). No significant differences were noted in the hazard of thromboembolism, bleeding, or death between the cohorts ($p>0.05$).

Conclusion. Patients with stable INRs while receiving warfarin who experience a significant subtherapeutic INR value have a low risk of thromboembolism in the ensuing 90 days. The risk was similar to that observed in a matched control population in whom therapeutic anticoagulation was maintained. These findings do not support the practice of anticoagulant bridge therapy for patients stabilized on warfarin therapy to reduce their risk for thromboembolism during isolated periods of subtherapeutic anticoagulation.

Key Words: vitamin K antagonist, warfarin, international normalized ratio, INR, thromboembolism, subtherapeutic anticoagulation, bridge therapy, anticoagulation management, thrombosis.

(Pharmacotherapy 2008;28(8):960–967)

Warfarin is highly effective for the prevention or treatment of both arterial and venous thromboembolism.[1–5] However, optimal use of warfarin is hampered by unpredictable pharmacokinetics; thus, even in specialized anticoagulant management clinics, about 40% of international normalized ratio (INR) values fall outside the target therapeutic range. Most out-of-range INR values are below the target range.[6–9] There are several reasons why patients may experience an unexpected INR below the usual therapeutic reference interval: dose omission, institution of alternate drugs that reduce the anticoagulant effect of warfarin, increases in dietary vitamin K intake, and otherwise unexplainable reductions in the anticoagulant effect of warfarin.[10–13] For warfarin-treated patients, thromboembolism has been associated with subtherapeutic INR values.[14] As a result, some clinicians prescribe bridge therapy with injectable anticoagulants, such as low-molecular-weight heparin (LMWH), in an attempt to reduce the risk of thrombosis.[15, 16]

Studies evaluating thromboembolic complications in patients receiving warfarin have generally associated risk of thrombosis with an INR value at the time of, or immediately preceding, the complicating event.[17, 18] Poor anticoagulation control has also been associated with increased risk of recurrent venous thromboembolism.[19] However, we found no studies that prospectively evaluated the absolute risk of thromboembolic complications among patients stabilized on anticoagulation therapy who during the normal course of therapy experience an INR significantly below their therapeutic range. To address this knowledge gap, we conducted a retrospective, longitudinal investigation to assess and quantify the absolute risk associated with isolated subtherapeutic anticoagulation in a large and carefully monitored cohort of patients receiving warfarin for a variety of indications.

From the Clinical Pharmacy Anticoagulation Service (Drs. Clark, Witt, and Trapp) and the Clinical Pharmacy Research Team (Dr. Delate), Kaiser Permanente Colorado, Lafayette, Colorado; the Department of Internal Medicine, University of New Mexico Health Sciences Center, Albuquerque, New Mexico (Dr. Garcia); the School of Clinical Medicine, University of Insubria, Varese, Italy (Dr. Ageno); the Department of Medicine, Research Unit–Section of General Internal Medicine, Boston University School of Medicine, Boston, Massachusetts (Dr. Hylek); and the School of Medicine, McMaster University, Hamilton, Ontario, Canada (Dr. Crowther).

Supported by the American College of Clinical Pharmacy Frontiers Fund: the 2006 Sanofi-Aventis Thrombosis Investigator Development Award.

Results previously reported in abstract form at the annual meeting of the American College of Clinical Pharmacy, Denver, Colorado, October 14–17, 2007.

Manuscript received November 15, 2007. Accepted pending revisions January 4, 2008. Accepted for publication in final form February 21, 2008.

Address reprint requests to Nathan P. Clark, Pharm.D., BCPS, CACP, Kaiser Permanente Colorado, 280 Exempla Circle, Lafayette, CO 80026; e-mail: nathan.clark@kp.org.

Methods

Study Design and Setting

This retrospective, longitudinal cohort analysis was conducted at Kaiser Permanente Colorado (KPCO), a health care delivery system providing integrated medical care to approximately 450,000 patients in the Denver-Boulder metropolitan area. All study activities were reviewed and approved by the KPCO institutional review board.

Anticoagulation services at KPCO are provided by a centralized Clinical Pharmacy Anticoagulation Service (CPAS).[6] This service provides comprehensive services for all KPCO patients requiring anticoagulation therapy and monitors more than 7000 patients. Working collaboratively with the referring physician, CPAS clinical pharmacists introduce anticoagulation therapy, order relevant laboratory tests, adjust anticoagulation drugs as necessary, and refill anticoagulation drug prescriptions. Details regarding operational aspects of this service have been described previously.[6] An integrated, electronic medical, pharmacy, and laboratory records system and the CPAS database (Dawn-AC; 4S Systems, Ltd., Cumbria, United Kingdom) were used to identify patients, treatments, and outcomes for this study.

Study Participants

Patients treated with warfarin during January 1, 1998 through December 31, 2005, with an INR value of 0.5 or more units below the patient-specific target INR range lower limit (defined as the "index low INR") were assigned to the low INR cohort. For example, an INR of 1.5 or lower would qualify as an index INR for a patient with a target INR range of 2.0–3.0. To identify newly low INR results, the two INR values preceding the index low INR had to be within or greater than the upper limit of the patient-specific INR target range and obtained at least 2 weeks apart.

Control patients (therapeutic INR cohort) with an INR value within the patient-specific therapeutic range (defined as the "index therapeutic INR") were identified during the

same time period. The two INR values preceding the index therapeutic INR had to be within or greater than the upper limit of the patient-specific INR target range and obtained at least 2 weeks apart. Patients in the therapeutic INR cohort could not have an INR measurement 0.2 or more units below the patient-specific target INR range lower limit at any time during the 90 days after the index INR. The INR criteria for cohort determination are graphically depicted in Figure 1.

Patients in the therapeutic INR cohort were matched to those in the low INR cohort at a ratio of up to 2:1 based on the index INR date (± 15 days), indication for warfarin therapy (venous thromboembolism, atrial fibrillation, heart valve disorder, or other), and age (± 5 yrs). All patients were required to have continuous enrollment in KPCO during the 180 days before the index INR. Patients were excluded if they were receiving warfarin with a target INR range lower limit of less than 2.0, or had a prescribed interruption of warfarin for any reason (e.g., an invasive procedure) during the 90-day follow-up. Patients in the low INR cohort could not be included in the therapeutic cohort during a later period of INR stability. The cohorts were mutually exclusive.

Study Outcomes and Data Collection

The primary outcome was occurrence of an anticoagulation-related thromboembolic complication during the 90 days after the index INR. The secondary outcomes were the times to the first occurrence of a verified anticoagulation therapy–related complication (thrombosis, hemorrhage, or death) during the 90 days after the index INR.

Thromboembolic complications were defined as any venous thromboembolism, cerebrovascular accident, transient ischemic attack, systemic embolism, or heart valve thrombosis. Bleeding complications included episodes such as gastrointestinal hemorrhage, hematoma, hemarthrosis, hemoptysis, and hematuria. Major hemorrhage was identified and defined as a bleeding event requiring transfusion of 2 or more units of red blood cells or surgical correction, an event causing a decrease in hemoglobin concentration of 2 g/dl or more, or any intracranial, intraarticular, intraocular, or retroperitoneal bleeding.

Complications were identified through queries of the KPCO integrated, electronic claims database by using predefined *International Classification of Diseases, Ninth Revision* (ICD-9) codes and verified through medical record review using a standardized data abstraction form. Deaths were identified from the KPCO integrated, electronic membership database. Fatal events were assessed for direct relationship to hemorrhage or thromboembolism and validated through review of the medical record and/or death certificate. The relationship to anticoagulation therapy for all identified complications was confirmed by two reviewers (N.C. and M.T.) blinded to cohort designation and the INR data at the time of the event, using a modified Naranjo adverse drug reaction probability scale.[20] Disagreements between reviewers were resolved by a third reviewer (M.C., E.H., or D.G.).

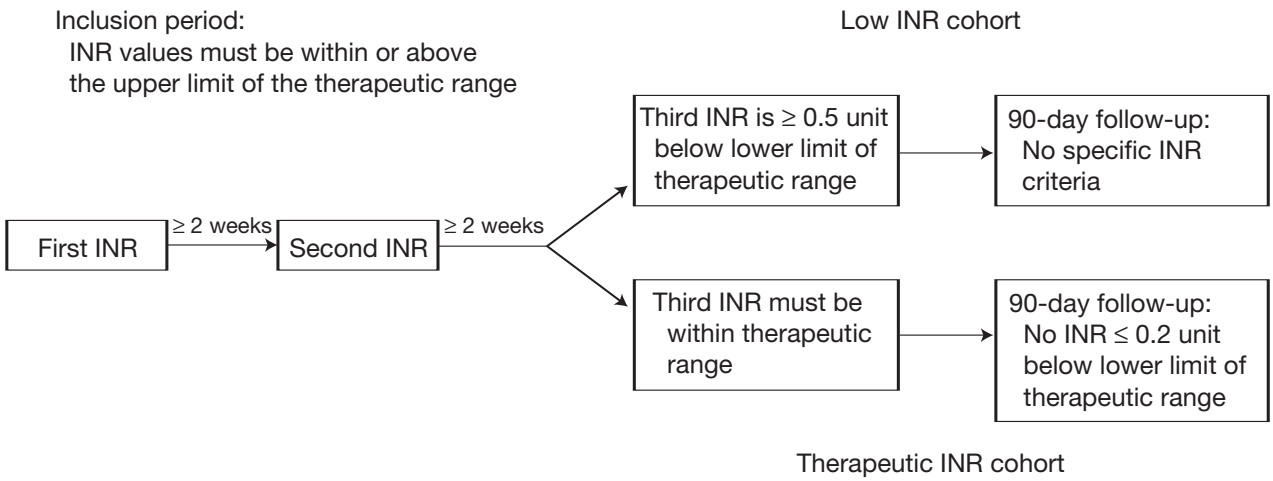


Figure 1. International normalized ratio (INR) requirements for designation of the low INR and therapeutic INR cohorts.

Information on comorbidities (hypertension, diabetes mellitus, heart failure, cancer [excluding squamous and basal cell carcinoma]) in the 180 days before the index INR and previous cerebrovascular accident, transient ischemic attack, thrombosis, and surgery in the 90 days before the index INR were identified from the KPCO clinic visits integrated, electronic database by using predefined ICD-9 codes. Information on age, sex, indication for warfarin use, target INR range, INR values, INR test dates, date of initiation of warfarin therapy, and warfarin dosing was obtained from the Dawn-AC database. Information on purchase of estrogen therapy (a thromboembolic risk factor) prescriptions during the 90 days preceding the index INR and use of heparin therapy in the 90 days after the index INR was obtained from the KPCO integrated, electronic pharmacy database.

Statistical Analysis

Baseline characteristics between patients in the low INR and therapeutic INR cohorts were compared by using the McNemar's test of association for proportions and the Wilcoxon test (for nonnormally distributed continuous variables) or matched *t* tests (for normally distributed continuous variables) for matched data. Conditional proportional hazards modeling was used to estimate the hazard ratios and their 95% confidence intervals (with and without adjustment for potential confounders) for anticoagulation therapy–related complications among the low INR cohort in relation to the therapeutic INR cohort. Censoring was performed at termination of KPCO membership or death, time of event, or 90 days after the index INR date, whichever came first. A unique match identification number was assigned to each matched set of patients and entered into the model as a cluster variable to account for the correlations within the matched sets and provide unbiased standard error estimates.

A global model was constructed with the dependent variable being the time interval from the index INR until the first occurrence of any anticoagulation therapy–related complication or censoring. In addition, conditional proportional hazards modeling was performed to estimate the hazard ratios and their 95% confidence intervals for each individual complication (bleeding, thromboembolism, or death). The models were constructed by using matching variables (index INR date, anticoagulation indication, and age) and variables that were statistically significantly different between the groups in the univariate analysis. Figure 2 contains information on the variables included in the models.

Results

A total of 5348 and 4110 patients met initial inclusion criteria for the therapeutic INR and low INR cohorts, respectively. After exclusions for noncontinuous KPCO enrollment and prescribed interruptions of warfarin therapy, 3469 and 1562 patients were available for matching in the therapeutic INR and low INR cohorts, respectively. A total of 1080 patients in the low INR cohort were matched to 1517 patients in the therapeutic INR cohort.

The indications for anticoagulation are listed in Table 1. The most common indication for anticoagulation was atrial fibrillation (46%). The proportions of patients with a diagnosis of

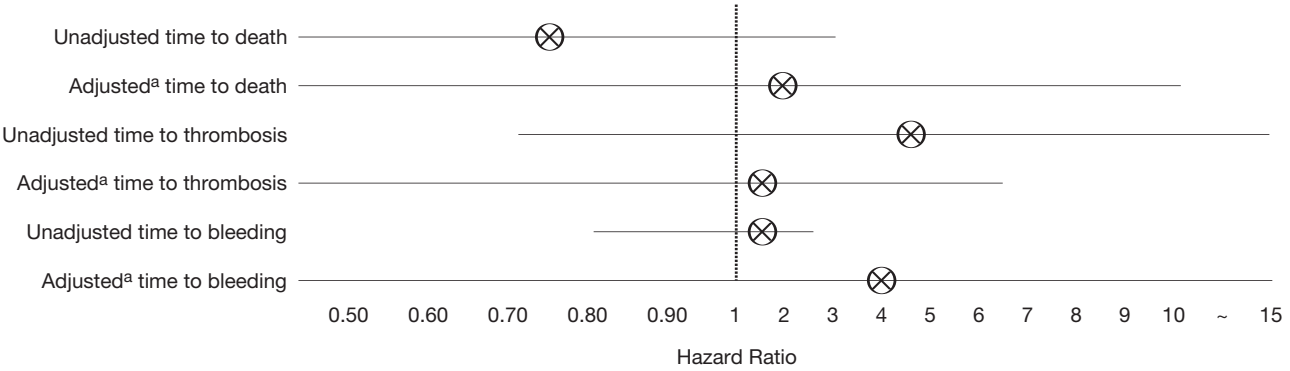


Figure 2. Hazard ratios for anticoagulation-related complications in the low INR cohort. [a]Adjusted for clustering of matched observations, matched variables, sex, percentage of previous INR values within the therapeutic range, diabetes mellitus, heart failure, and previous thrombosis.

Table 1. Baseline Characteristics by INR Status

Characteristic	Low INR Cohort (n=1080)	Therapeutic INR Cohort (n=1517)	p Value[a]
	Mean ± SD		
Age (yrs)[b]	68.3 ± 13.5	68.5 ± 13.0	0.472
In-range INRs during pre–index INR period (%)	65.1 ± 30.0	55.6 ± 26.1	<0.001
	No. (%) of Patients		
Male	534 (49.4)	816 (53.8)	0.018
INR target			
2.5	998 (92.4)	1459 (96.2)	<0.001
≥ 3.0	82 (7.6)	58 (3.8)	<0.001
Primary indication for anticoagulation therapy			
Atrial fibrillation	494 (45.7)	693 (45.7)	0.977
Venous thromboembolism	372 (34.4)	546 (36.0)	0.416
Heart valve disorder	46 (4.3)	53 (3.5)	0.361
Other	168 (15.6)	225 (14.8)	0.580
Thromboembolic risk factors			
Diabetes mellitus	222 (20.6)	255 (16.8)	0.053
Hypertension[c]	346 (32.0)	520 (34.3)	0.158
Heart failure[c]	180 (16.7)	221 (14.6)	0.165
History of thrombosis[c, d]	211 (19.5)	384 (25.3)	<0.001
History of stroke[c]	1 (0.1)	3 (0.2)	0.603
Cancer[c]	8 (0.7)	17 (1.1)	0.236
Recent surgery[e]	22 (2.0)	29 (1.9)	0.939
Estrogen therapy[e]	87 (8.1)	118 (7.8)	0.637
Thrombophilia[f]	45 (4.2)	65 (4.3)	0.878
Missed warfarin dose before index INR	298 (27.6)	8 (0.5)	<0.001
	Median (interquartile range)		
Index INR	1.4 (1.3–1.5)	2.6 (2.3–3.1)	<0.001
No. of days to index INR measurement[g]	458 (151–1305)	162 (70–370)	<0.001

INR = international normalized ratio.
[a]Adjusted for clustering of matched observations and matching variables (if variable was used in matching, adjusted for clustering and other matching variables).
[b]As of date of index INR measurement.
[c]During the 180 days before the index INR.
[d]Any arterial or venous thrombosis.
[e]During the 90 days before the index INR.
[f]Anticardiolipin antibody, antiphospholipid antibody, antithrombin deficiency, elevated factor VIII activity, factor V Leiden, hyperhomocysteinemia, lupus anticoagulant, nephrotic syndrome, protein C deficiency, protein S deficiency, prothrombin 20210 mutation.
[g]From start of warfarin therapy.

diabetes, hypertension, heart failure, previous stroke, and cancer were similar between the cohorts (p>0.05). The therapeutic INR cohort had a higher proportion of male patients, patients with an INR target of 2.5, and patients with a history of thrombosis.

The overall rate of anticoagulation-related thromboembolic complications was low and similar in both groups (p=0.214; Table 2). There were five anticoagulation-related thromboembolic events, four (0.4%) in the low INR cohort and one (0.1%) in the therapeutic INR cohort. These included three transient ischemic attacks, one cerebrovascular accident, and one upper extremity arterial thromboembolism. The INR value was below the therapeutic range at the time of thrombosis in three of the events and was not available in the remaining two events. There were no fatal thromboembolic events. Table 3 presents the details of these events.

No statistically significant differences were noted between the cohorts in proportions of patients with anticoagulation-related bleeding events (p=0.151) or death (p=0.766; Table 2). Also, no statistically significant differences were noted in the hazard ratios of anticoagulation-

Table 2. Characteristics by INR Status During the 90-Day Follow-up Period

Characteristic	Low INR Cohort (n=1080)	Therapeutic INR Cohort (n=1517)	p Value[a]
	Mean ± SD		
In-range INRs during follow-up (%)	51.0 ± 30.7	68.2 ± 33.8	<0.001
Percent change in warfarin dose	7.4 ± 17.4	-0.9 ± 5.4	<0.001
	No. (%) of Patients		
Received heparin or LMWH	13 (1.2)	6 (0.4)	0.075
Received warfarin dose boost[b]	702 (65.0)	5 (0.3)	<0.001
Deceased	37 (3.4)	48 (3.2)	0.689
Anticoagulation-related complications			
Death	2 (0.2)	3 (0.2)	0.766
Thrombosis	4 (0.4)	1 (0.1)	0.214
Bleeding	16 (1.5)	12 (0.8)	0.151

INR = international normalized ratio; LMWH = low-molecular-weight heparin.
[a]Adjusted for clustering of matched observations and matching variables.
[b]A 1- or 2-day dose escalation (1.25–2 x usual daily dose) in response to index INR.

Table 3. Details of the Five Thromboembolic Events During the 90-Day Follow-up

Indication for Anticoagulation	Event	Patient Characteristics	Description
Arterial TE	Arterial TE	83-year-old woman with a history of heart failure, diabetes mellitus, hyperlipidemia	Embolus to left third and fourth fingers, improved with unfractionated heparin and nifedipine
CVA	TIA	70-year-old man with a history of several cerebral ischemic events, hypertension, diabetes mellitus	Slurred speech and dizziness, CT scan negative, symptoms resolved spontaneously
Atrial fibrillation	CVA	83-year-old woman with history of heart failure, hypertension, and pulmonary embolism	Right-hemisphere CVA
Prosthetic mitral valve	TIA	48-year-old woman with a St. Jude's bileaflet valve, atrial fibrillation, and diabetes mellitus	Numbness and weakness of left upper extremity, CT scan negative, symptoms resolved
CVA	TIA	80-year-old man with history of carotid stenosis, diabetes mellitus, hypertension, and coronary artery disease	Acute mental status changes and incontinence, CT scan negative twice

TE = thromboembolism; TIA = transient ischemic attack; CT = computed tomography; CVA = cerebral vascular accident.

related thrombosis, bleeding, or death between the cohorts for both the unadjusted and adjusted time-to-event analysis in the 90 days after the index INR (Figure 2).

Ten major hemorrhagic events occurred: three (0.3%) in the low INR cohort and seven (0.5%) in the therapeutic INR cohort (p>0.05). These included five fatal bleeding events (four intracranial hemorrhages and one gastrointestinal bleed): two in the low INR cohort and three in the therapeutic INR cohort (p>0.05). The mean ± SD INR at the time of bleeding event was 3.2 ± 1.9 in the low INR cohort and 2.9 ± 0.6 in the therapeutic INR cohort. Overall, deaths were similar in both groups, occurring in 3.4% of the low INR cohort and 3.2% of the therapeutic INR cohort (p=0.689).

Patients in the therapeutic INR cohort had a higher proportion of out-of-range INR values in the pre–index INR period compared with those patients in the low INR cohort (44% vs 35%, p<0.001). In the 90-day follow-up period, the therapeutic INR cohort had a higher proportion of INR values in range (p<0.001; Table 2). Of note, 28% of patients in the low INR cohort had documentation of missed warfarin doses before the index INR (Table 1).

Sixty-five percent of patients in the low INR cohort received a warfarin dose boost, defined as a 25–100% increase in the daily dose for 1 or 2 days in response to the index low INR. The mean weekly warfarin dose increased by 7% in

response to the index INR in the low INR cohort and was unchanged in the therapeutic INR cohort (Table 2). The first INR value after the index INR remained 0.5 unit below the therapeutic INR range in 138 patients (12.7%) in the low INR cohort, and 58 patients (5.4%) had a third consecutive subtherapeutic INR. None of the four thrombotic events occurred among these subgroups.

Discussion

This retrospective study of established warfarin patients with relatively stable INR control enrolled in a large anticoagulation service found no association between an isolated low INR and increased risk for objectively confirmed arterial or venous thromboembolism during the 90-day follow-up period. The risk of thromboembolism in these patients was similar to that seen in matched control patients without similar low INRs during the 90-day follow-up. This finding suggests that the absolute risk associated with an unexpected isolated subtherapeutic INR is low.

Subtherapeutic anticoagulation has been established as a risk factor for stroke in atrial fibrillation.[14, 21] In a previous case-control study, the odds of stroke doubled with INR values equal to 1.7 and tripled with an INR of 1.5 or less, compared with INR values of 2.0 or greater.[14] A comprehensive review of anticoagulant therapy for patients with mechanical heart valves similarly suggests increased risk of valve thrombosis and arterial thromboembolism with lower intensity anticoagulation.[22] Patients with unprovoked venous thromboembolism treated for 3 months with standard intensity anticoagulation (INR 2.0–3.0) who were then randomly assigned to lower intensity anticoagulation (INR 1.5–1.9) had increased risk of recurrent venous thromboembolism compared with those continuing with standard anticoagulation intensity.[23] These data suggest subtherapeutic INR values confer an increased risk for subsequent thrombosis. However, to our knowledge, no previous study has attempted to prospectively quantify the 90-day absolute risk of thrombosis associated with isolated subtherapeutic INR values that occur commonly in routine clinical practice.

It is important to consider the setting within which patients were managed when interpreting our results. Timely and careful follow-up by CPAS pharmacists to ensure that the INR returned to the target range quickly may have contributed to the low event rate. These results may be less generalizable to institutions where timely follow-up is not the standard. In addition, the effect of a warfarin dose boost is unknown but may have facilitated a quicker return to the therapeutic range. Unfractionated heparin and LMWH use during the follow-up period was infrequent (1.2%), reducing the likelihood that bridge therapy produced the observed low rates of thromboembolism.

Given the low rate of thromboembolic events, we interpret our results to suggest that patients stabilized on warfarin therapy who present with a single low INR value do not require any specific therapy other than warfarin dosage adjustment and increased frequency of INR monitoring to ensure a rapid return to their prescribed reference INR range. Our observations do not support the use of bridge therapy with a rapid-acting, injectable anticoagulant such as LMWH for these patients. Such therapy is unlikely to reduce the rate of thrombosis significantly, is expensive, and has the potential to cause bleeding.[24]

Strengths of this study include its large sample size, carefully defined patient populations, uniform and complete follow-up and outcome ascertainment, independent review of low INRs to rule out prescribed warfarin interruption, and blinded independent adjudication of outcome events. The retrospective nature of the study may have resulted in incomplete capture of some outcome events. However, the rigorous monitoring provided by CPAS minimized the likelihood of undetected clinically important outcome events such as bleeding, thrombosis, or death. Our findings were reported at the population level, and extrapolation to high-risk subgroups, such as patients with mechanical heart valves and recent (within 1 mo) venous thromboembolic events, should be done with caution. The hazard ratio estimates in the time-to-event analysis should be interpreted cautiously because of the small number of events observed. Nevertheless, our study included a large sample, increasing the likelihood that clinically important differences in the outcome event rate between the two groups enrolled in the study should have been detected.

Conclusion

Patients stabilized on warfarin therapy who experience an episode of subtherapeutic anticoagulation (INR ≥ 0.5 unit below the lower limit of their target range) have a low risk for

thromboembolism in the ensuing 90 days. The risk was similar to that observed in a matched control population in which therapeutic anticoagulation was maintained. These findings do not support the practice of bridge therapy with a rapid-acting parenteral anticoagulant for patients stabilized on warfarin therapy in order to reduce their risk of thromboembolism during isolated periods of significant subtherapeutic anticoagulation.

References

1. **Hutten BA, Prins MH**. Duration of treatment with vitamin K antagonists in symptomatic venous thromboembolism. Cochrane Database Syst Rev 2006;(1):CD001367.
2. **Anand SS, Yusuf S**. Oral anticoagulant therapy in patients with coronary artery disease: a meta-analysis. JAMA 1999;282:2058–67.
3. **Meade TW, Wilkes HC, Kelleher CC, et al**. 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 high risk. Lancet 1998;351:233–41.
4. **van Walraven C, Hart RG, Singer DE, et al**. Oral anticoagulants vs aspirin in nonvalvular atrial fibrillation: an individual patient meta-analysis. JAMA 2002;288:2441–8.
5. **Schulman S, Granqvist S, Holmstrom M, et al**. The duration of oral anticoagulant therapy after a second episode of venous thromboembolism. N Engl J Med 1997;336:393–8.
6. **Witt DM, Sadler MA, Shanahan RL, Mazzoli G, Tillman DJ**. Effect of a centralized clinical pharmacy anticoagulation service on the outcomes of anticoagulant therapy. Chest 2005;127: 1515–22.
7. **Gurwitz JH, Field TS, Radford MJ, et al**. The safety of warfarin therapy in the nursing home setting. Am J Med 2007;120: 539–44.
8. **Reynolds MW, Fahrbach K, Hauch O, et al**. Warfarin anticoagulation and outcomes in patients with atrial fibrillation: a systematic review and meta-analysis. Chest 2004;126:1938–45.
9. **Hylek EM, Evans-Molina C, Shea C, Henault LE, Regan S**. Major hemorrhage and tolerability of warfarin in the first year of therapy among elderly patients with atrial fibrillation. Circulation 2007;115:2689–96.
10. **Booth SL, Charnley JM, Sadowski JA, et al**. Dietary vitamin K1 and stability of oral anticoagulation: proposal of a diet with constant vitamin K1 content. Thromb Haemost 1997;77:504–9.
11. **O'Reilly RA, Rytand DA**. "Resistance" to warfarin due to unrecognized vitamin K supplementation. N Engl J Med 1980;303:160–1.
12. **Holbrook AM, Pereira JA, Labiris R, et al**. Systematic overview of warfarin and its drug and food interactions. Arch Intern Med 2005;165:1095–106.
13. **Cropp JS, Bussey HI**. A review of enzyme induction of warfarin metabolism with recommendations for patient management. Pharmacotherapy 1997;17:917–28.
14. **Hylek EM, Skates S, Sheehan M, Singer DE**. An analysis of the lowest effective intensity of prophylactic anticoagulation for patients with nonrheumatic atrial fibrillation. N Engl J Med 1996;335:540–6.
15. **Kovacs MJ, Kearon C, Rodger M, et al**. Single-arm study of bridging therapy with low-molecular-weight heparin for patients at risk of arterial embolism who require temporary interruption of warfarin. Circulation 2004;110:1658–63.
16. **Douketis JD, Johnson JA, Turpie AG**. Low-molecular-weight heparin as bridging anticoagulation during interruption of warfarin. Arch Int Med 2004;164:1319–26.
17. **Palareti G, Manotti C, D'Angelo A, et al**. Thrombotic events during oral anticoagulant treatment: results of the inception-cohort, prospective, collaborative ISCOAT study. Thromb Haemost 1997;78:1438–43.
18. **Oake N, Fergusson DA, Forster AJ, van Walraven C**. Frequency of adverse events in patients with poor anticoagulation: a meta-analysis. Can Med Assoc J 2007;176:1589–94.
19. **Palareti G, Legnani C, Cosmi B, et al**. Poor anticoagulation quality in the first 3 months after unprovoked venous thromboembolism is a risk factor for long-term recurrence. J Thromb Haemost 2005;3:955–61.
20. **Naranjo CA, Busto U, Sellers EM, et al**. A method for estimating the probability of adverse drug reactions. Clin Pharmacol Ther 1981;30:239–45.
21. **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.
22. **Salem DN, Stein PD, Al-Ahmad A, et al**. Antithrombotic therapy in valvular heart disease: native and prosthetic. The seventh ACCP conference on antithrombotic and thrombolytic therapy. Chest 2004;1236:S457–82.
23. **Kearon C, Ginsberg JS, Kovac MJ, et al**. Comparison of low-intensity warfarin therapy with conventional-intensity warfarin therapy for long-term prevention of recurrent thromboembolism. N Engl J Med 2003;349:631–9.
24. **Tillman DJ, Charland SL, Witt DM**. Effectiveness and economic impact associated with a program for outpatient management of acute deep vein thrombosis in a group model health maintenance organization. Arch Intern Med 2000;160:2926–32.

BRIEF REPORT

Results of a Medication Reconciliation Survey from the 2006 Society of Hospital Medicine National Meeting

Brian J. Clay, MD[1]
Lakshmi Halasyamani, MD[2]
Erin R. Stucky, MD[3]
Jeffrey L. Greenwald, MD[4]
Mark V. Williams, MD[5]

[1] Assistant Professor of Medicine, Division of Hospital Medicine, University of California San Diego, California.

[2] Vice Chair, Department of Internal Medicine, Saint Joseph Mercy Hospital, Ann Arbor, Michigan.

[3] Clinical Professor of Pediatrics, University of California San Diego; Pediatric Hospital Medicine Fellowship Director, Rady Children's Hospital San Diego, California.

[4] Director, Hospital Medicine Unit, Boston Medical Center; Associate Professor of Medicine, Boston University School of Medicine, Boston, Massachusetts.

[5] Chief, Division of Hospital Medicine, Northwestern University Feinberg School of Medicine, Chicago, Illinois.

BACKGROUND: The status of implementation of medication reconciliation across hospitals is variable to date; the degree to which hospitalists are involved is not known.

METHODS: To better describe the current state of medication reconciliation implementation, we conducted a survey of attendees of the 2006 Society of Hospital Medicine national meeting.

RESULTS: We identified a lack of uniformity across hospitals with respect to the degree of process implementation. Hospitalists were involved in design and implementation in a majority of cases, and felt that medication reconciliation would likely have a positive impact on patient safety. Tertiary care academic centers were more likely to use physicians to perform medication reconciliation, whereas community hospitals were more likely to involve nurses as well. Pharmacist participation in the medication reconciliation process was found to be quite low. Process and outcome measures were used infrequently. Patients' lack of medication knowledge and absence of preadmission medication information were cited most frequently as barriers to implementation of medication reconciliation.

CONCLUSIONS: Implementation of medication reconciliation is complex and challenging. Medication information is often incomplete, and elements of the medication reconciliation process result in increased time demands on providers. Current implementation efforts often have physicians and nurses "share" responsibility for compliance, and pharmacists are underutilized in medication reconciliation processes. Hospitalists have thus far played a substantial role in process design and implementation, and should continue to lead the way in advancing efforts to successfully implement medication reconciliation. ***Journal of Hospital Medicine* 2008;3(6):465–472.**

KEYWORDS: medication reconciliation, patient safety, quality control, Society of Hospital Medicine.

The Joint Commission's (TJC) National Patient Safety Goal (NPSG) #8—"Accurately and completely reconcile medications across the continuum of care"—challenges hospitals to design and implement new medication management processes. With medication errors contributing to patient morbidity and mortality,[1] establishing a comprehensive process for reconciling a patient's medications during the hospitalization episode is an important quality improvement and patient safety goal.

However, the current state of inpatient medication management is highly fragmented. Standard documentation is lacking, as is integration of information between care settings.[2] There are now reports describing implementation of various medication reconciliation processes for admissions,[3] transfers,[4] and discharges.[5]

DOI 10.1002/jhm.370
Published online in Wiley InterScience (www.interscience.wiley.com).

Hospitalists are well-positioned to contribute to the implementation of medication reconciliation. Indeed, because TJC does not explicitly specify what type of health care provider (eg, physician, nurse, etc.) should assume responsibility for this process, institutions have designed workflows to suit their own needs, while striving to comply with national standards.

Given the complexity and lack of standardization around this NPSG, a survey was distributed to attendees of a Society of Hospital Medicine (SHM) national meeting to determine the various processes implemented thus far, and to ascertain existing challenges to implementation. We report here on the results.

METHODS

A survey tool (Appendix) was designed to query demographic and institutional factors, involvement in the process, and barriers to implementation of medication reconciliation. Surveys were included in all attendees' registration materials, resulting in the distributions of approximately 800 surveys.

Responses were entered into an Excel spreadsheet. Simple descriptive statistics were used to determine proportions for providers, processes, and barriers to implementation. Where appropriate, variables were dichotomized, allowing for paired *t*-test analysis. Statistical significance was defined as a *P* value less than .05. Subgroup analyses by hospital type, provider type, and process method were performed.

RESULTS

A total of 295 completed surveys were collected. The responses are tabulated in Table 1.

TABLE 1
Survey Responses

Primary practice setting	
Academic tertiary center	23%
Community teaching hospital	29%
Non-academic hospital	43%
Patient population	
Adults only	90%
Pediatrics only	5%
Adults and pediatrics	5%
State of implementation	
Fully implemented	48%
Partially implemented	35%
Planning stages	11%
Unaware of plans to implement	2%
Unaware of med reconciliation	4%
Hospitalist involvement	
Active role	36%
Peripheral role	24%
No role	31%
Process format	
Paper	47%
Computer	11%
Both paper and computer	31%
Don't know	2%
Measuring compliance	
Yes	42%
No	14%
Don't know	34%
Measuring outcomes	
Yes	22%
No	25%
Don't know	41%
Impact of medication reconciliation	
No impact	9%
Positive impact	58%
Negative impact	7%
Don't know	14%

Response totals may not always total 100% due to some answers being left blank. Percentages reported are of the total of 295 surveys.

Process

A paper process was used most often (47%), followed by a combined process (31%), and computers alone in just 11% of cases. Measurement of process compliance was reported in less than half (42%), with 34% unaware if their institutions were monitoring compliance. Outcome measurement was recorded as "not performed" (25%) or "unknown" (41%) in a majority of cases. Respondents reported a favorable view of the future impact of medication reconciliation, with 58% citing likely positive impacts on patient safety and patient care; fewer were unsure (14%) or anticipated no impact (9%) or negative impact (7%). Survey results regarding responsibility for individual process steps are detailed in Table 2. Notably, respondents often indicated that both physicians and nurses would share responsibility for a given step. Physicians were more often responsible for reconciling home medications, updating discharge medication lists, and communicating to outpatient providers. Nursing performed reconciliation in only 10% of cases. Results across all steps demonstrated very low participation rates by pharmacists, with pharmacist responsibility for reconciliation only 6% of the time.

Hospital Type

Results of subgroup analyses by hospital type are detailed in Table 3. Community teaching hospitals

TABLE 2
Survey Responses – Medication Reconciliation Process Steps

Process Step	Physician	Nurse	Physician and Nurse	Pharmacist	Other
Obtaining home med list	15%	39%	41%	3%	2%
Documenting home med list	17%	41%	37%	2%	3%
Reconciling medications	56%	10%	21%	6%	7%
Updating discharge med list	64%	6%	17%	3%	10%
Providing instructions at discharge	15%	46%	32%	2%	5%
Communicating changes at follow-up	84%	6%	4%	6%	1%

Response totals may not always total 100% due to some answers being left blank. Percentages reported are of the total of 295 surveys.

(CTHs) were significantly more likely (57%) than nonteaching hospitals (NTHs) (49%) or tertiary academic centers (TACs) (35%) to have achieved full implementation. NTHs were significantly less likely to have involved hospitalists in implementation. Use of computer-based processes at TACs was more common (27%) than in CTHs (9%) or NTHs (7%). TACs were significantly more likely to have a physician obtain the medication list (33%, compared with 15% and 7% for CTHs and NTHs, respectively), whereas NTHs were more likely to use nurses (50%) than were CTHs (31%) or TACs (26%). Similar significant differences were found among hospital types with regard to obtaining the preadmission medication list. Physicians in TACs (25%) were more likely to be responsible for giving discharge medication instructions than in CTHs (10%) or NTHs (14%, not significant compared with TACs).

Barriers

Results regarding barriers to successful implementation are shown in Table 4. Patient lack of knowledge of medications (87%) and absence of a preadmission medication list from other sources (80%) were common. Both paper and computer medication reconciliation processes were associated with respondents citing cumbersome hospital systems as a barrier; this barrier was cited more often when the implemented process was paper-only (Table 5). Respondents who stated the medication reconciliation process takes too long did so regardless of whether the implemented process was paper-based or computer-based. Despite these barriers, only 16% of respondents stated that medication reconciliation was not worth the effort of implementation. Barriers reported were similar across hospital type (Table 6) with 2 exceptions. Formulary differences were noted to be a barrier more often in CTHs (78%) compared with NTHs (60%) and TACs (64%, not significant compared with CTHs). Language barriers were problematic more often in TACs (48%) than in NTHs (28%) or CTHs (36%, not significant compared with TACs).

DISCUSSION

Managing medication information for inpatients is an extremely complex task. On admission, home medication lists are often inaccurate or absent,[6] requiring extra time and effort to discover this information. By discharge, medication regimens have frequently been altered,[7] making communication of changes to the next provider essential. One study described myriad provider, patient, and health system issues in maintaining accurate outpatient medication lists.[8] These issues are further compounded by the multiple prescribers, necessary hand-offs, and formulary differences in the inpatient setting.

Over half of the hospitalists in this survey reported hospitalist involvement in design and implementation of medication reconciliation. Given the familiarity with hospital systems and inpatient workflow, hospitalists are well-positioned to contribute to successful implementation. Nonetheless, many were unaware of efforts to implement this NPSG.

Measurement of both process and outcome measures is important when determining value in quality improvement. Beyond process measures, outcome measures such as adverse drug events, readmission rates, mortality, patient satisfaction, and outpatient provider satisfaction may be appropriate in evaluating medication reconciliation strategies. Even measuring the accuracy of the process with respect to the admission orders written would be a valuable source of information for further improvement. Unfortunately, respondents indicated that evaluation was occurring infrequently. Potentially more problematic is the apparent lack of

TABLE 3
Subgroup Analysis by Hospital Type

	Academic Centers [AC]	Community Teaching Hospitals [CT]	Non-Teaching Hospitals [NT]	*P* values (2-tailed)		
				AC vs. CT	AC vs. NT	CT vs. NT
State of implementation						
Fully implemented	25/71 (35)	48/84 (57)	68/139 (49)	0.007	0.06	0.25
Partially implemented	31/71 (44)	25/84 (30)	48/139 (35)	0.07	0.21	0.44
Planning stages	9/71 (13)	9/84 (11)	14/139 (10)	0.70	0.51	0.81
Unaware of plans to implement	2/71 (3)	1/84 (1)	3/139 (2)	0.37	0.65	0.57
Unaware of med reconciliation	4/71 (5)	1/84 (1)	6/139 (4)	0.14	0.74	0.19
Hospitalist involvement						
Active role	28/59 (47)	34/80 (43)	43/127 (34)	0.64	0.09	0.19
Peripheral role	12/59 (20)	25/80 (31)	34/127 (27)	0.15	0.30	0.54
No role	19/59 (32)	19/80 (24)	50/127 (39)	0.30	0.36	0.03
Process format						
Paper	26/59 (44)	47/81 (58)	63/127 (50)	0.10	0.45	0.26
Computer	16/59 (27)	7/81 (9)	9/127 (7)	0.005	<0.001	0.60
Both paper and computer	17/59 (29)	25/81 (31)	51/127 (40)	0.80	0.15	0.19
Don't know	0/59 (0)	2/81 (2)	4/127 (3)	0.28	0.18	0.66
Process steps (selected questions)						
Obtaining home med list						
Physician	19/58 (33)	12/80 (15)	9/125 (7)	0.013	<0.001	0.07
Physician and Nurse	19/58 (33)	39/80 (49)	49/125 (39)	0.47	0.44	0.16
Nurse	15/58 (26)	25/80 (31)	62/125 (50)	0.005	0.003	0.008
Pharmacist	5/58 (9)	1/80 (1)	2/125 (2)	0.06	0.03	0.58
Documenting home med list						
Physician	22/58 (38)	11/80 (14)	11/125 (9)	0.001	<0.001	0.26
Physician and Nurse	15/58 (26)	37/80 (46)	45/125 (36)	0.02	0.18	0.16
Nurse	18/58 (31)	26/80 (32)	64/125 (51)	0.90	0.012	0.008
Pharmacist	3/58 (5)	2/80 (3)	1/125 (1)	0.55	0.09	0.29
Reconciling medications						
Physician	33/58 (57)	51/80 (64)	63/125 (50)	0.41	0.42	0.051
Physician and Nurse	8/58 (14)	14/80 (18)	32/125 (26)	0.53	0.09	0.18
Nurse	6/58 (10)	6/80 (8)	15/125 (12)	0.68	0.71	0.36
Pharmacist	8/58 (14)	5/80 (6)	3/125 (2)	0.11	0.007	0.13
Updating discharge med list						
Physician	42/58 (72)	50/80 (63)	76/125 (61)	0.27	0.15	0.77
Physician and Nurse	7/58 (12)	16/80 (20)	23/125 (18)	0.22	0.31	0.72
Nurse	2/58 (3)	5/80 (6)	10/125 (8)	0.41	0.20	0.59
Pharmacist	3/58 (5)	3/80 (4)	3/125 (2)	0.78	0.27	0.40
Providing instructions at discharge						
Physician	14/57 (25)	8/80 (10)	17/125 (14)	0.02	0.07	0.40
Physician and Nurse	14/57 (25)	30/80 (38)	39/125 (31)	0.11	0.41	0.30
Nurse	25/57 (44)	37/80 (46)	60/125 (48)	0.82	0.62	0.80
Pharmacist	4/57 (7)	1/80 (1)	0/125 (0)	0.06	0.003	0.26

Results are tabulated only out of those surveys with answers for the particular question. Percentage results are listed in parentheses.
Response totals may not always total 100% due to some respondents entering an answer of "Other."

clarity regarding identification of healthcare provider responsibility for specific process steps. By far the least uniformity is in the acquisition and documentation of the preadmission medication list. There is variability in who is assigned to perform this task, but a substantial number of respondents indicated that their process involved a "shared" responsibility between physicians and nurses. It is unclear whether this phenomenon reflects the complexity of inpatient medication information management, or is simply an attempt to distribute the work among providers. Sharing the work between physicians and nurses may increase the overall likelihood for compliance and possibly improve the safety and accuracy of the process, especially if the physicians and nurses take the medication history

in a redundant fashion and share their findings. Conversely, compliance may decrease if each provider merely expects the other to complete the process. Optimally, an interdisciplinary workflow for medication history taking would be in place, involving both physicians and nurses, with the availability of pharmacist consultation in complex cases. However, our survey data suggest this is infrequent; resident physicians appear to be the ones shouldering substantial responsibility for medication reconciliation in tertiary academic centers. Further research into the accuracy of medication reconciliation processes involving different strategies for medication information collection would be useful.

We documented several barriers to successful implementation of medication reconciliation. Physicians cited a lack of medication knowledge on the part of the patient and unavailable prior medication lists as substantial barriers to success. Many medication reconciliation processes are limited by issues of poor health literacy or inadequate patient knowledge about medications. This lack of medication knowledge is especially problematic for patients new to a healthcare system. It will be important to implement processes that not only reconcile medications accurately, but also make medication information available for future care episodes.

Time required to complete the process was also important. Certain elements of the medication reconciliation process are "new work," and integrating the process into existing workflows is crucial. Given the significant time commitment required, the rare involvement of pharmacists at most institutions is striking. It appears that hospital pharmacists do not currently "own" any of the medication reconciliation process steps at most facilities, despite having formal training in medication history-taking. In the 2006 ASHP national hospital pharmacy survey, one-third of pharmacists stated that there were not enough pharmacy resources to meet medication reconciliation demands; only 19% of those surveyed stated pharmacists provided medication education at discharge to more than 25% of their patients.[9]

This report has several limitations. The survey used was not comprehensive, and only represents a convenience sample of hospitalists attending a national meeting. Nearly 300 physicians

TABLE 4
Survey Results – Barriers to Implementation

Barrier to Implementation	Yes	No	Unsure
Patient not knowing meds	87%	2%	0%
Process takes too long	53%	28%	8%
Med list not available	80%	9%	0%
Process not worth effort	16%	60%	12%
Cumbersome hospital systems	52%	33%	4%
Formulary differences	59%	24%	5%
Language barriers	31%	53%	4%
No access to outside records	63%	23%	2%
Lack of job clarity in process	38%	48%	3%
Availability of med list at discharge	27%	57%	3%

Response totals may not always total 100% due to some answers being left blank. Percentages reported are of the total of 295 surveys.

TABLE 5
Subgroup Analysis of Barriers to Implementation by Process Type

Barriers (Selected Questions)	Paper Only [P]	Computer Only [C]	Paper and Computer [PC]	*P* values (2-tailed)		
				P vs. C	P vs. PC	C vs. PC
Process takes too long						
Yes	77/134 (57)	19/31 (61)	55/91 (60)	0.69	0.65	0.92
No	43/134 (32)	11/31 (35)	28/91 (31)	0.75	0.87	0.68
Unsure	14/134 (10)	1/31 (3)	8/91 (9)	0.21	0.80	0.27
Process not worth effort						
Yes	24/133 (18)	3/31 (10)	17/91 (19)	0.28	0.85	0.25
No	93/133 (70)	22/31 (71)	62/91 (68)	0.91	0.75	0.76
Unsure	16/133 (12)	6/31 (19)	12/91 (13)	0.30	0.82	0.41
Cumbersome hospital systems						
Yes	86/133 (65)	16/31 (52)	46/92 (50)	0.18	0.03	0.85
No	42/133 (32)	13/31 (42)	42/92 (46)	0.29	0.03	0.70
Unsure	5/133 (4)	2/31 (6)	4/92 (4)	0.62	0.82	0.64

Results are tabulated only out of those surveys with answers for the particular question. Percentage results are listed in parentheses.
Response totals may not always total 100% due to rounding.

TABLE 6
Subgroup Analysis of Barriers to Implementation by Hospital Type

Barrier to Implementation (Selected Questions)	Academic Centers [AC]	Community Teaching Hospitals [CT]	Non-Teaching Hospitals [NT]	*P* values		
				AC vs. CT	AC vs. NT	CT vs. NT
Process takes too long						
Yes	37/58 (64)	49/78 (63)	70/124 (56)	0.90	0.31	0.37
No	15/58 (26)	24/78 (31)	42/124 (34)	0.53	0.28	0.66
Unsure	6/58 (10)	5/78 (6)	12/124 (10)	0.39	0.88	0.32
Process not worth effort						
Yes	7/58 (12)	16/78 (21)	23/123 (19)	0.17	0.24	0.73
No	42/58 (72)	52/78 (67)	84/123 (68)	0.53	0.59	0.88
Unsure	9/58 (16)	10/78 (12)	16/123 (13)	0.50	0.59	0.84
Cumbersome hospital systems						
Yes	36/58 (62)	46/79 (58)	69/123 (56)	0.64	0.45	0.78
No	19/58 (33)	32/79 (41)	46/123 (37)	0.34	0.60	0.57
Unsure	3/58 (5)	1/79 (1)	8/123 (7)	0.16	0.61	0.049
Formulary differences						
Yes	37/58 (64)	61/78 (78)	74/123 (60)	0.07	0.61	0.009
No	16/58 (28)	14/78 (18)	41/123 (33)	0.17	0.50	0.02
Unsure	5/58 (8)	2/78 (3)	8/123 (7)	0.19	0.81	0.22
Language barriers						
Yes	28/58 (48)	28/77 (36)	34/123 (28)	0.16	0.009	0.24
No	28/58 (48)	46/77 (60)	82/123 (67)	0.17	0.016	0.32
Unsure	2/58 (3)	3/77 (4)	7/123 (5)	0.76	0.54	0.74
No access to outside records						
Yes	38/58 (66)	60/79 (76)	87/123 (71)	0.20	0.50	0.44
No	18/58 (31)	18/79 (23)	33/123 (27)	0.30	0.58	0.52
Unsure	2/58 (3)	1/79 (1)	3/123 (2)	0.39	0.68	0.58
Lack of job clarity in process						
Yes	26/58 (45)	31/79 (39)	49/121 (40)	0.48	0.53	0.89
No	28/58 (48)	46/79 (58)	68/121 (56)	0.25	0.32	0.78
Unsure	4/58 (7)	2/79 (3)	4/121 (3)	0.28	0.22	0.75
Availability of med list at discharge						
Yes	20/58 (34)	24/79 (30)	35/120 (29)	0.62	0.50	0.88
No	36/58 (62)	54/79 (68)	78/120 (65)	0.47	0.70	0.66
Unsure	0/58 (0)	1/79 (1)	7/120 (6)	0.45	0.06	0.08

Results are tabulated only out of those surveys with answers for the particular question. Percentage results are listed in parentheses.
Response totals may not always total 100% due to rounding.

responded, representing both teaching and private hospital settings. We consider the response rate of 37% reasonable for a survey of this nature, and the variety of processes described is likely indicative of the overall status of medication reconciliation implementation. The over-representation of certain institutions in our survey is possible, especially those with large or influential hospital medicine programs. Our survey did not ask respondents to name their home institutions. In addition, this design is open to a convenience sample bias, in that surveying only national meeting attendees (rather than the entire SHM membership) risks overinclusion of those hospitalists involved in leadership roles and quality improvement projects. Despite this, the variety of processes described is likely indicative of the overall status of medication reconciliation implementation in mid-2006. It is possible that processes have become more uniform nationwide in the interim.

Our survey results reflect the complexity surrounding medication reconciliation. It appears that full implementation has not yet occurred everywhere, significant barriers remain, and outcome measurement is limited. Importantly, physicians, nurses, and pharmacists do not have standardized roles. Responsibility for medication reconciliation has predominantly been added to the existing duties of inpatient physicians and nurses, with limited involvement of pharmacists. Hospitalists are well-

positioned to lead the ongoing implementation of medication reconciliation processes and should take advantage of their systems knowledge to effectively partner with other physicians, nurses, and pharmacists to achieve success in medication reconciliation.

ACKNOWLEDGMENTS

The authors thank Ken Epstein, MD, and Renee Meadows, MD, along with the entire SHM Medication Reconciliation Task Force for their helpful review and comments on the article.

APPENDIX

2006 SHM National Meeting Medication Reconciliation Survey Questions

1. Which setting best describes your primary practice setting?
   Academic, tertiary-care center — Community-based academic center
   Non-academic community hospital — Other:_____

2. For which patient population do you provide care?
   Adults only — Pediatrics only — Adults and pediatrics

3. Mark the single best statement that describes the current state of medication reconciliation implementation at your institution:

   Fully implemented across all departments
   Implemented in some hospital departments but not all
   Still in planning stages of how to implement medication reconciliation
   I am unaware of any plans to implement medication reconciliation
   Have never heard of medication reconciliation

4. Were you or another hospitalist that you are aware of involved in the development or implementation of the medication reconciliation process at your organization?
   Yes – played an active role (e.g. member of committee or task force charged with development and implementation)
   Yes – played a peripheral role (reviewed process and implementation plan once developed)
   No

5. Which process best describes your medication reconciliation process
   Paper-based — Both paper- and computer-based components
   Computer-based — Don't know

6. Are you measuring compliance with your process in any way? Yes No Don't know

7. Are you measuring outcomes of your process in any way? Yes No Don't know

8. How do you think your current medication reconciliation process will affect patient care and patient safety?
   No impact — Negative impact
   Positive impact — Don't know

9. Please mark the person(s) who are performing each of listed aspects of medication reconciliation:
   P = physician; N = nurse; Ph = pharmacist.

9.1	Obtaining the current list of patient home medications	☐ P ☐ N ☐ Ph ☐ P and N ☐ Other
9.2	Documenting list of patient home medications	☐ P ☐ N ☐ Ph ☐ P and N ☐ Other
9.3	Reconciling home medications with currently ordered	☐ P ☐ N ☐ Ph ☐ P and N ☐ Other
9.4	Updating/documenting changes to patient's discharge	☐ P ☐ N ☐ Ph ☐ P and N ☐ Other
9.5	Providing medication instructions to the patient	☐ P ☐ N ☐ Ph ☐ P and N ☐ Other
9.6	Communicating medication changes to follow-up clinicians	☐ P ☐ N ☐ Ph ☐ P and N ☐ Other

10. Please identify issues from the list below that are potential barriers in your current medication reconciliation process:

10.1	Patient does not know medications:	☐ Yes	☐ No	☐ Unsure
10.2	Med recon process takes too long:	☐ Yes	☐ No	☐ Unsure
10.3	Medication list not available:	☐ Yes	☐ No	☐ Unsure
10.4	Med recon process not worth the effort:	☐ Yes	☐ No	☐ Unsure
10.5	Cumbersome hospital systems:	☐ Yes	☐ No	☐ Unsure
10.6	Inpatient/outpatient formulary differences:	☐ Yes	☐ No	☐ Unsure
10.7	Language barrier:	☐ Yes	☐ No	☐ Unsure
10.8	Difficulty getting outside records:	☐ Yes	☐ No	☐ Unsure
10.9	Lack of job clarity on who does med	☐ Yes	☐ No	☐ Unsure
10.10	Availability of med list on discharge:	☐ Yes	☐ No	☐ Unsure

Address for correspondence and reprint requests: Dr. Brian Clay, University of California San Diego, 200 West Arbor Drive, Mail Code 8485, San Diego, CA 92103. Telephone: 619-471-9194; Fax: 619-543-8255; E-mail: bclay@ucsd.edu

Received 5 November 2007; revision received 2 February 2008; accepted 14 March 2008.

REFERENCES

1. Kohn LT, Corrigan JM, Donaldson MS, eds. *To Err Is Human: Building a Safer Health System*. Washington, DC: National Academy Press; 1999.
2. Barnsteiner JH. Medication reconciliation: transfer of medication information across settings – keeping it free from error. *Am J Nurs*. 2005;105(3 Suppl):31–36.
3. Gleason KM, Groszek JM, Sullivan C, et al. Reconciliation of discrepancies in medication histories and admission orders of newly hospitalized patients. *Am J Health-Syst Pharm*. 2004;61(16):1689–1695.
4. Pronovost P, Weast B, Schwartz M, et al. Medication reconciliation: a practical tool to reduce the risk of medication errors. *J Crit Care*. 2003;18(4):201–205.
5. Paquette-Lamontagne N, McLean WM, Besse L, Cusson J. Evaluation of a new integrated discharge prescription form. *Ann Pharmacother*. 2001;35(7-8):953–958.
6. Rozich J, Howard RJ, Justeson JM, et al. Patient safety standardization as a mechanism to improve safety in health care. *Jt Comm J Qual Saf*. 2004;30(1):5–14.
7. Himmel W, Tabache M, Kochen MM. What happens to long-term medication when general practice patients are referred to hospital? *Eur J Clin Pharmacol*. 1996;50(4):253–257.
8. Varkey P, Reller MK, Smith A, et al. An experiential interdisciplinary quality improvement education initiative. *Am J Med Qual*. 2006;21(5):317–322.
9. Pedersen CA, Schneider PJ, Scheckelhoff DJ. ASHP national survey of pharmacy practice in hospital settings: monitoring and patient education-2006. *Am J Health-Syst Pharm*. 2007;64(5):507–520.

Journal of Physical Activity and Health, 2008, 5, 469-480

Psychosocial Correlates of Exercise in Women With Self-Reported Depressive Symptoms

Lynette L. Craft, Frank M. Perna, Karen M. Freund, and Larry Culpepper

Background: Exercise effectively reduces symptoms of depression. However, correlates of regular exercise in depressed women are unknown. This study assessed psychosocial determinants of exercise in a sample of women with depressive symptoms. ***Methods:*** Sixty-one women completed demographic, depression, and exercise-related questionnaires. ***Results:*** The average Primary Health Questionnaire-9 (PHQ-9) depression score was 12.1 (SD = 5.0), indicating moderate depressive symptoms. In the previous week, the women reported 12.8 metabolic equivalents (METs) of exercise. Low levels of self-efficacy and social support for exercise were also reported. Depressive symptoms were positively associated with barriers to exercise ($r = .35$, $P < .01$), and barriers were inversely related to exercise METs ($r = -.37$, $P < .05$). Barriers to activity and education level were significant determinants of exercise. ***Conclusions:*** Women with depressive symptoms report minimal exercise involvement, numerous barriers to exercise, and low exercise self-efficacy and social support for exercise.

Keywords: physical activity, mental health, women's health

It is estimated that roughly 18 million adults or 9.5% of the US population age 18 years and older suffer from clinical depression.[1] Depression is also a risk factor for several chronic diseases such as cardiovascular disease, diabetes, and obesity, further contributing to the overall health impact of this disorder.[2-5] Depression is twice as prevalent in women as men, with the lifetime prevalence of major depression estimated as high as 23% in middle-aged women.[6] In primary care settings, approximately 5% to 28% of patients suffer from major or minor depressive disorders.[6] Although many individuals do not meet diagnostic criteria for clinical depression, subclinical depressive symptoms can also negatively impact physical health

Craft is with the Dept of Preventive Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611. Perna and Freund are with the Section of General Internal Medicine and Women's Health, Interdisciplinary Research Center, and Culpepper the Dept of Family Medicine, Boston University School of Medicine, Boston, MA 02118.

and are associated with morbidity.[7] Consequently, depression (even at subclinical levels) signifies a threat to the psychological and physical well-being of millions of women and is an area of women's health that warrants further examination.

Exercise has been shown to be an effective adjunct to traditional treatments for depression. It is as effective as psychological therapies for depression and might be as effective as antidepressant medication for some.[8-11] Several meta-analyses conducted on this literature confirm that exercise involvement is associated with a reduction in the symptoms of depression, that exercise is effective for many types of patients, and that the exercise session need not be lengthy or intense to produce positive effects.[12-15] Overall effect sizes from these meta-analyses range from –0.72 to –1.4, indicating moderate to large effects.

Research suggests that long-term exercise programs (ie, at least 9 weeks) and programs promoting current public health recommendations for physical activity are more effective in alleviating symptoms of depression than exercise programs of shorter duration.[12,16] Therefore, exercise interventions will be most effective for those who are able to maintain exercise involvement on a regular basis (ie, at least 3 to 5 days a week). Cross-sectional research in this field suggests that depressed patients are more sedentary and have reduced physical work capacities as compared with nondepressed patients.[17,18] However, little is known about the psychosocial determinants of exercise among women reporting depressive symptoms.

In the general population, fatigue, lack of time, lack of an exercise partner, self-consciousness, and care-giving issues are commonly sited as reasons for not exercising regularly.[19,20] Women with depressive symptoms might face barriers to exercise that are similar to nondepressed women; however, it is also possible that they perceive additional barriers specific to their illness (ie, psychomotor retardation, lack of motivation, etc). Likewise, in healthy samples, addressing key factors specific to the individual (eg, exercise self-efficacy, social support for exercise, perceived barriers to exercise) can also increase rates of adherence to exercise programs by up to 25%.[21-23] Previous research supports the use of cognitive behavioral strategies such as self-monitoring, goal setting, and feedback to improve exercise adherence.[23-25] Of note, perhaps the most important factor predictive of exercise behavior is brief, but frequent, supportive contact with the participant.[23,24,26,27]

We aimed to ascertain the degree to which commonly endorsed exercise determinants are relevant for women reporting depressive symptoms. With a better understanding of the determinants of exercise in this group, strategies to promote exercise involvement and improve exercise adherence can be developed and implemented. A better understanding of barriers to and promoters of exercise involvement in this group should, in turn, improve the likelihood that exercise can be effectively incorporated as a beneficial adjunct treatment for clinical depression. Therefore, the purpose of the current study was to gather initial descriptive data regarding exercise determinants from a diverse sample of women with depressive symptoms. It was hypothesized that depressive symptoms would be associated with the number of perceived barriers to exercise. Furthermore, we hypothesized that barriers would be inversely associated with exercise involvement. Finally, we also anticipated that barriers to exercise, social support for exercise, and exercise self-efficacy would be significantly associated with exercise involvement in this group.

Method

Participants

Before data collection, all participants signed an informed consent statement as approved by the Institutional Review Board at Boston University School of Medicine. Women between the ages of 18 and 65 years, self-identifying as experiencing depressive symptoms, were recruited for this study. To be eligible for participation, women had to score 5 (indicating mild depressive symptoms) or greater on the Primary Health Questionnaire-9 (PHQ-9) and could be currently sedentary or physically active. Women were excluded from participation if they self-reported the presence of a comorbid mental health diagnosis (eg, anxiety disorder, bipolar disorder, schizophrenia) or any physical health problem that prevented them from exercising. Women were recruited from the greater Boston area and were referred to the study by primary care physicians, psychologists, or responded to posted study flyers. Women from primary care clinics in a community hospital (Boston Medical Center) were specifically targeted for enrollment in an attempt to obtain a sample of women that was representative of women reporting depressive symptoms in primary care settings and more racially and socioeconomically inclusive than previous exercise and depression research. Each woman received $5.00 for her participation. Sixty-one women were included in the study, and sample characteristics are presented in Table 1.

Questionnaires

Demographic Questionnaire. All participants completed a basic demographic questionnaire that assessed variables such as age, marital status, race, family income level, medical history, comorbid mental health diagnoses, and current treatment for depression.

The Primary Health Questionnaire-9 (PHQ-9). Presence of depressive symptoms was confirmed using the PHQ-9. The PHQ-9 is a reliable and valid self-administered 9-item depression module based on DSM-IV criteria and is commonly used in primary care settings to screen for depression.[28] Participants are asked to indicate how often in the past 2 weeks they had been bothered by each of the 9 DSM-IV criteria-based symptoms. Items are ranked from 0 (*not at all*) to 3 (*nearly every day*). Scores of 5 to 9 represent mild, 10 to 14 moderate, 15 to 19 moderately severe, and 20 or greater severe depression.[29] A cut-score of 10 is used to signify clinically relevant symptoms. Compared with a mental health professional interview, the PHQ-9 has demonstrated good sensitivity (88%) and specificity (88%) in the diagnosis of major depression, and in the general population, it has been found valid in the diagnosis of major depression and its subtypes and sensitive to changes in depression over time.[28] Furthermore, it is highly correlated with other self-report measures used to assess depression such as the Beck Depression Inventory[30] and is appropriate for use with racially and ethnically diverse women including Caucasians, African Americans, and Latinas.[31] Internal consistency for this sample was satisfactory, with Cronbach $\alpha = .76$.

Table 1 Sample Characteristics (N = 61)

Characteristic	%
Age	mean = 38.6 y (SD = 9.7)
Ethnic background	
Caucasian	50.0
African American	45.0
Latina	5.0
Marital status	
single	51.7
married/partnered	26.7
separated/divorced	21.6
Employment	
full-time	49.2
part-time	16.4
not employed	34.4
Income	
<20,000	56.7
21,000–30,000	11.7
31,000–55,000	13.3
56,000–75,000	8.3
>75,000	10.0
Education	
not a high school graduate	6.6
high school graduate	19.7
some college	34.4
college graduate	31.1
some graduate school	8.2
Percentage of participants with mild, moderate, moderately severe, and severe symptoms of depression	
PHQ-9 Score 5–9	33.3
PHQ-9 Score 10–14	41.7
PHQ-9 Score 15–19	15.0
PHQ-9 Score ≥20	10.0
Current treatment for depressive symptoms	
medication	50.8
psychotherapy	31.7

Godin Leisure Time Exercise Questionnaire (LTEQ). The LTEQ is a reliable and widely used self-report instrument that surveys frequency of mild, moderate, and vigorous leisure-time physical activity for the past 7 days.[32] The questionnaire yields number of bouts per week of light, moderate, and strenuous activity. Work output can be estimated by assigning metabolic equivalent values (3, 5, 9 METs) to each exercise category and multiplying by frequency of occurrence. Total METs in the previous week was used as the outcome of interest.

The Perceived Barriers to Physical Activity Scale (PBPA). Perceived barriers were assessed with an 8-item questionnaire that has been found to predict exercise behavior in general epidemiologic investigations.[33] Items were rated on a 5-point Likert scale (*never* to *very often*) to reflect the frequency with which a perceived barrier (eg, no one to exercise with, other people discourage me, feeling too tired) interfered with or prevented a respondent from exercising. Furthermore, we calculated the total number of barriers that were rated a 3 or higher (indicating that the barrier sometimes, often, or very often interfered) for each participant.

The Social Support Scale for Exercise Behavior (SSEB). This inventory has 29 items and measures family and friend support for exercise, which has been associated with exercise involvement.[34] Three subscales are generated from the questionnaire: anticipated family and friend involvement, family rewards and punishments, and friends exercising together. For each item, participants are asked to rate from 1 (*never*) to 5 (*very often*) how frequently a supportive behavior occurred in the previous 3 months. Examples of supportive behaviors are "offered to exercise with me," "discussed exercise with me," and "gave me encouragement to stick with my exercise program." Scores are summed, and higher scores are indicative of higher social support. Cronbach alphas for the 3 subscales in this sample were: anticipated family and friend involvement, $\alpha = .81$; friends exercising together, $\alpha = .92$; and family rewards and punishments, $\alpha = .62$. Because of the low reliability of the family rewards and punishments subscale, it was not used in statistical analyses.

Self-Efficacy for Exercise Behavior (SEEB). The SEEB is a 14-item questionnaire that assesses 2 factors associated with exercise maintenance: making time for exercise and resisting relapse.[29] Participants are asked to rate on a scale from 1 (*sure I could not do it*) to 5 (*sure I could do it*) their confidence in their ability to motivate themselves to do a variety of exercise behaviors consistently for at least 6 months. Examples of items include "stick to your exercise program when you have excessive demands at work," "exercise even though you are feeling depressed," and "stick to your exercise program even when the weather is bad." Scores are summed, and higher scores indicate higher levels of exercise self-efficacy. The 2 subscales demonstrated acceptable internal consistency in this sample: making time for exercise, $\alpha = .85$, and resisting relapse, $\alpha = .91$.

Procedure

Potential participants met with a member of the research staff and signed the informed consent statement. Next, the PHQ-9 was completed. Women scoring 5 or greater were next asked to complete the demographic questionnaire, which included questions regarding the presence of comorbid mental health diagnoses and significant physical health problems. Finally, for those deemed eligible for participation (ie, scoring 5 or greater on the PHQ-9 and indicating no comorbid mental health diagnoses or physical health problems contraindicating exercise), the remaining exercise-related questionnaires were completed. The research staff member was present to assist the participant with questionnaire completion and answer any questions that arose.

Analyses

Descriptive statistics were computed for study variables. Pearson correlations were also calculated to examine relationships among exercise (METs), depressive symptoms, and psychosocial variables. Finally, linear regression analyses were conducted to identify significant determinants of exercise after controlling for baseline depression score, age, and family income.

Results

Seventy-five women met the initial eligibility criteria of scoring 5 or greater on the PHQ-9. However, 11 women were excluded from participation because of the self-report of a comorbid mental health diagnosis [generalized anxiety disorder (n = 5), posttraumatic stress disorder (n = 4), bipolar disorder (n = 2)]. In addition, 3 participants were exercising substantially more (>3 SD above the mean) than the rest of the sample. Therefore, these 3 participants were also removed from the statistical analyses, resulting in a final sample size of 61 women.

The PHQ-9 depression scores in this sample ranged from 5 to 23, with an average score of 12.1 (SD = 5.0), indicative of moderate depressive symptoms. Approximately one-third (33%) of the women were experiencing subclinical levels of depressive symptoms (ie, PHQ-9 score < 10). In the previous week, the women reported an average of only 0.28 bouts (SD = 0.99) of strenuous activity, 1.6 bouts (SD = 1.4) of moderate activity, and 0.82 bouts (SD = 1.9) of mild activity of at least 15 minutes in length in their leisure time. The average METs reported for the previous week were 12.8 (SD = 15.6, median = 10.0). Forty-eight percent were exercising 30 minutes or less per week, and 84.4% were not meeting the American College of Sports Medicine's current physical activity recommendations.[35]

Correlations among study variables were examined to determine relationships among depression, exercise, and psychosocial determinants. No significant correlations were found between depressive symptoms and number of bouts of moderate and vigorous activity ($r = -.12$, $P = .41$) or between depressive symptoms and total METs expended ($r = -.11$, $P = .46$). However, there was a significant correlation between severity of depression and number of barriers to exercise ($r = .35$, $P < .01$). Furthermore, number of perceived barriers to exercise and total METs expended were inversely correlated ($r = -.37$, $P < .05$). In addition, number of perceived barriers to exercise was inversely correlated with exercise self-efficacy (resisting relapse: $r = -.28$, $P < .05$; making time: $r = -.36$, $P < .01$). Thus, the data suggest that perceived barriers might play an important role in the relationship between depressive symptoms and activity involvement. Correlations are presented in Table 2.

The most frequently reported barriers to exercise were feeling too tired (84.1%), a lack of time (61.9%), feeling self-conscious about one's looks when exercising (54.0%), and not having an exercise partner (52.4%). Seventy percent of the women indicated that they had 3 or more barriers that sometimes, often, or very often interfered with their attempts to exercise. When types of perceived barriers were examined relative to severity of depressive symptoms, a similar pattern emerged (see Table 3). The women in this study also indicated lower self-efficacy to make time for exercise (sample mean = 23.9, SD = 8.0 versus norm mean = 28.2, SD = 6.7) and lower friend social support (sample mean = 9.6, SD = 8.3 versus norm

Table 2 Correlations Among Exercise, Depressive Symptoms, and Psychosocial Variables

	1	2	3	4	5	6	7
1. METs/wk		−.11	−.37[a]	.09	.10	00	−.08
2. Depression score (PHQ-9)			.35[b]	−.15	−.21	−.04	.16
3. Number of barriers to exercise				−.28[a]	−.36[b]	.06	.18
4. Exercise self-efficacy–resisting relapse					.78[b]	.29[a]	−.06
5. Exercise self-efficacy–making time						.22	−.17
6. Social support–family involvement							.49[b]
7. Social support–friend involvement							—

[a] Correlation is significant at the .05 level.

[b] Correlation is significant at the .01 level.

Table 3 Most Frequently Reported Barriers to Exercise by Symptom Severity

Barrier	Percentage of participants
Mild (N = 20)	
feeling too tired	80%
not enough time	65%
lack of exercise partner	50%
feeling self-conscious	50%
Moderate (N = 26)	
feeling too tired	96%
not enough time	68%
lack of exercise partner	56%
feeling self-conscious	56%
Moderately severe (N = 9)	
feeling too tired	75%
lack of exercise partner	75%
not enough time	63%
feeling self-conscious	63%
Severe (N = 10)	
feeling too tired	100%
feeling self-conscious	83%
lack of safe place to exercise	66%
not enough time	50%

mean = 13.6, SD = 5.3) and family social support (sample mean = 21.8, SD = 15.3 versus norm mean = 26.1, SD = 10.9) as compared with published norms.[29,34]

Finally, linear regression analyses were conducted to identify significant determinants of time spent in physical activity. After controlling for factors known to be associated with both depression and leisure-time activity (age, severity of depressive symptoms, and family income), number of barriers to exercise ($\beta = -0.31$, $P = .05$), education level ($\beta = 0.43$, $P = .01$), and length of current depressive episode ($\beta = -0.32$, $P = .07$) were the best determinants of exercise (total METs), accounting for 28.5% of the variance.

Discussion

This study is novel in the attempt to gather information regarding psychosocial correlates of exercise among a sample of women with self-reported symptoms of depression. Our current understanding of exercise determinants results from research conducted in the general population, with healthy samples. Consequently, less is known regarding psychosocial correlates of exercise among women with depressive symptoms. Such information is extremely important given the current interest in using exercise as an adjunct treatment for depressive symptoms.

Consistent with previous cross-sectional research on depressed women and epidemiological research of physical activity patterns in the United States,[17,18,36] the vast majority of participants were sedentary. Exercise recommendations from physicians, therapists, and other mental health clinicians are likely to go a long way toward impacting behavior change. In primary care settings, brief exercise counseling advice, as well as telephone contact, is associated with a reduction in perceived barriers to exercise and improvements in fitness and exercise adherence.[37,38]

The women in this study endorsed barriers to exercise that are quite similar to those reported by the general population, with fatigue and a lack of time the most commonly cited barriers. Moreover, in support of our hypothesized relationships between depressive symptoms, barriers, and exercise, most women in the study reported more than 3 barriers that interfere with exercise attempts, and greater depressive symptoms were associated with more perceived barriers to exercise. Furthermore, the number of perceived barriers and length of current depressive symptoms were indicative of time spent in exercise.

Perceptions of self-efficacy and social support for exercise were also low in this group. That is, participants felt little confidence to engage in regular exercise and identified few instances of support for exercise attempts from either family members or friends. These findings might reflect negative evaluations that are commonly activated among individuals with depression. That is, depressed individuals are often subject to a negative perceptual bias in which they interpret situations and social interactions in a manner that is consistent with their own perception of themselves, the world, and their future as negative and bleak.[39] Future research is needed to clarify whether perceptions of low social support for exercise are the result of such a perceptual bias or an actual lack of social support (which might also be a consequence of depression).

These are important findings in light of research from the general population, which consistently implicates barriers, self-efficacy, and social support as predictive of exercise involvement.[21-23,25,40-42] Not surprisingly, women with more severe

symptoms or those who have been experiencing symptoms for a greater length of time might find exercise particularly difficult to initiate in light of the barriers they perceive. Thus, these findings highlight the importance of assessing and addressing these psychosocial constructs (barriers, efficacy, and social support) when encouraging women with depressive symptoms to begin exercise programs. Similarly, exercise interventions aimed at reducing symptoms of depression should incorporate exercise counseling that focuses on identifying and reducing barriers to exercise. Furthermore, interventions should promote a supportive mastery climate and cognitive restructuring to enhance feelings of self-efficacy and social support.

There are several limitations to the current study that should be considered. First, depressive symptoms were self-reported and then verified by a widely used and validated screening tool. Although depression is commonly identified in this way in larger epidemiological investigations, it remains preferable to have a trained clinician confirm a formal diagnosis of depression. Consequently, this lack of a formal mental health diagnosis represents a limitation of this investigation, and findings might not generalize to women diagnosed with clinical depression. However, in light of the impact that subclinical depression can have on both mental and physical health, we believe that our sample represents an important cross-section of women seeking treatment in primary care settings and, therefore, contributes to our understanding of the relationships among exercise, depressive symptoms, and psychosocial correlates.

Most of the minority women who participated in the study were of African American descent. As such, several other racial/ethnic minority groups are underrepresented in this sample. Social and cultural factors might differ among various ethnic groups and, as a result, have a differing impact on depressive symptoms and involvement in regular physical activity. The small sample used in this study did not allow us to analyze the data for minority women separately. Thus, more research is needed to examine the relationships among these variables in minority populations. In addition, the cross-sectional nature of the data prevents us from making causal associations. The data represent 1 sample of women with depressive symptoms at one time point. The relationships examined might change over time or during the course of a depressive episode. For example, it is impossible to determine whether depressive symptoms cause patients to perceive more barriers to exercise or if barriers to exercise contribute to women feeling more depressed. Although the women did report a range of severity of depressive symptoms, the average level of depressive symptoms was moderate. For women with more severe symptoms, it might be difficult to initiate exercise until traditional treatments (ie, medication) have had an effect. Therefore, the current findings might best characterize the attitudes and beliefs toward exercise of women who are experiencing mild to moderate symptoms of depression.

We employed commonly used measures of exercise determinants and found theoretically expected relationships between these psychosocial determinants and exercise, but women with depressive symptoms might have distinct barriers and attitudes toward exercise that are not tapped by these general measures. Elicitation studies with depressed women might yield measures that are more relevant and more strongly associated with exercise than general measures. Furthermore, we did not directly compare the women in this study to a group of nondepressed women. As such, it is unclear whether the findings we report are specific to women with

depressive symptoms. Although we have tried to make comparisons to the general population or nondepressed women when possible, we acknowledge that this is a limitation of our study, and additional research that directly compares depressed to nondepressed women is needed. Finally, relatively few of the women in the sample (~15.6%) were meeting current public health recommendations for physical activity. As a result, it is difficult to understand the ways in which women who are able to initiate and maintain exercise programs, in spite of their depressive symptoms, differ from the sedentary women in this sample. More research is needed to expose the correlates and predictors of activity among women with depressive symptoms adhering to exercise programs. Such information would be invaluable in counseling depressed women regarding exercise behavior change.

In summary, this study represents the first attempt to identify determinants of exercise in a diverse group of women with self-reported depressive symptoms. Our findings imply that, consistent with psychosocial determinants of activity in the general population, there are important relationships between perceived barriers, social support, exercise self-efficacy, and actual time spent engaged in exercise in this sample. Consequently, these particular determinants of exercise should be discussed when encouraging patients with depressive symptoms to initiate and maintain programs of regular exercise, keeping in mind that patients might need assistance in identifying barriers as well as solutions for reducing or eliminating barriers and enhancing self-efficacy and social support.

Acknowledgments

This work was supported by the National Institutes of Health: Clinical Research Feasibility Funds (M01RR00533) and the Office of Women's Health Research (K-12-HD4244).

References

1. Greenberg PE, Kessler RC, Birnbaum HG, et al. The economic burden of depression in the United States: how did it change between 1990 and 2000? *J Clin Psychiatry.* 2003;64:1465–1475.
2. Arroyo C, Hu FB, Ryan LM, et al. Depressive symptoms and risk of type 2 diabetes in women. *Diabetes Care.* 2004;27:129–133.
3. Carnethon MR, Kinder LS, Fair JM, Stafford RS, Fortmann SP. Symptoms of depression as a risk factor for incident diabetes: findings from the National Health and Nutrition Examination Epidemiologic Follow-Up Study, 1971–1992. *Am J Epidemiol.* 2003;158:416–423.
4. Chapman DP, Perry GS, Strine TW. The vital link between chronic disease and depressive disorders. *Prev Chronic Dis.* 2005;2(1). http://www.cdc.gov/pcd/issues/2005/jan/04_0066.htm. Accessed April 2007.
5. Ferketich AK, Frid DJ. Depression and coronary heart disease: a review of the literature. *Clin Geriatr.* 2001;9:1–8.
6. Bromberger JT. A psychosocial understanding of depression in women: for the primary care physician. *J Am Med Womens Assoc.* 2004;59:198–206.
7. Wells KB, Stewart A, Hays RD, et al. The functioning and well-being of depressed patients: results from the Medical Outcomes Study. *JAMA.* 1989;262:914–919.
8. Blumenthal JA, Babyak MA, Moore KA, et al. Effects of exercise training on older patients with major depression. *Arch Intern Med.* 1999;159:2349–2356.

9. Klein MH, Greist JH, Gurman AS, et al. A comparative outcome study of group psychotherapy vs. exercise treatments for depression. *Int J Ment Health*. 1985;13:148–177.
10. McNeil JK, LeBlanc EM, Joyner M. The effect of exercise on depressive symptoms in the moderately depressed elderly. *Psychol Aging*. 1991;6:487–488.
11. Veale D, LeFevre K, Pantelis C, et al. Aerobic exercise in the adjunctive treatment of depression: a randomized controlled trial. *J R Soc Med*. 1992;85:541–544.
12. Craft LL, Landers DM. The effect of exercise on clinical depression and depression resulting from mental illness: a meta-analysis. *J Sport Exerc Psychol*. 1998;20:339–357.
13. North TC, McCullagh P, Tran ZV. Effect of exercise on depression. *Exerc Sport Sci Rev*. 1990;18:379–415.
14. Lawlor DA, Hopker SW. The effectiveness of exercise as an intervention in the management of depression: systematic review and meta-regression analysis of randomized controlled trials. *BMJ*. 2001;322:763–767.
15. Stathopoulou G, Powers MB, Berry AC, Smits AJ, Otto MW. Exercise interventions for mental health: a quantitative and qualitative review. *Clin Psychol Sci Pract*. 2006;13:179–193.
16. Dunn AL, Trivedi MH, Kampert JB, Clark CG, Chambliss HO. Exercise treatment for depression: efficacy and dose response. *Am J Prev Med*. 2005;28:1–8.
17. Martinsen EW, Strand J, Paulsson G. Physical fitness levels in patients with anxiety and depressive disorders. *Int J Sports Med*. 1989;10:58–61.
18. Morgan WP. A pilot investigation of physical working capacity in depressed and nondepressed males. *Res Q Exerc Sport*. 1969;40:859–861.
19. Brownson RC, Eyler AA, King AC, et al. Patterns and correlates of physical activity among US women 40 years and older. *Am J Public Health*. 2000;90:264–270.
20. King AC, Castro C, Wilcox S, et al. Personal and environmental factors associated with physical inactivity among different racial-ethnic groups of US middle-aged and older-aged women. *Health Psychol*. 2000;19:354–364.
21. Buckworth J, Wallace LS. Application of the Transtheoretical Model to physically active adults. *J Sports Med Phys Fitness*. 2002;42:360–367.
22. Duncan KA, Pozehl B. Staying on course: the effects of an adherence facilitation intervention on home exercise participation. *Prog Cardiovasc Nurs*. 2002;17:59–65.
23. King AL, Taylor CB, Haskell WL, et al. Strategies for increasing early adherence to and long-term maintenance of home-based exercise training in healthy middle-aged men and women. *Am J Cardiol*. 1988;61:628–632.
24. Lombard DN, Lombard TN, Winett RA. Walking to meet health guidelines: the effect of prompting frequency and prompt structure. *Health Psychol*. 1995;14:164–170.
25. Wilbur J, Vassalo A, Chandler P, McDevitt J, Miller AM. Midlife women's adherence to home-based walking during maintenance. *Nurs Res*. 2005;54:33–40.
26. Lee RE, Goldberg JH, Sallis JF, Hickmann SA, Castro CM, Chen AH. A prospective analysis of the relationship between walking and mood in sedentary ethnic minority women. *Women Health*. 2001;32:1–15.
27. Writing Group for the Activity Counseling Trial Research Group. Effects of physical activity counseling in primary care: the ACT: a randomized controlled trial. *JAMA*. 2001;286:677–687.
28. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*. 2001;16:606–613.
29. Sallis JF, Pinski RB, Grossman RM, et al. The development of self-efficacy scales for health-related diet and exercise behaviors. *Health Educ Res*. 1988;3:283–292.
30. Martin A, Rief W, Klaiberg A, Braehler E. Validity of the brief patient health questionnaire mood scale (PHQ-9) in the general population. *Gen Hosp Psychiatry*. 2006;28:71–77.

31. Huang FY, Chung H, Kroenke K, Dulucchi KL, Spitzer RL. Using the patient health questionnaire-9 to measure depression among racially and ethnically diverse primary care patients. *J Gen Intern Med.* 2006;21:547–552.
32. Godin G, Shephard RJ. A simple method to assess exercise behavior in the community. *Can J Appl Physiol.* 1995;10:141–146.
33. Sallis JF, Hovell MF, Hofstetter CR. Predictors of adoption and maintenance of vigorous physical activity in men and women. *Prev Med.* 1992;21:237–251.
34. Sallis JF, Grossman RM, Pinski RB, et al. The development of scales to measure social support for diet and exercise behaviors. *Prev Med.* 1987;16:825–836.
35. Haskell WL, Lee IM, Pate RR, et al. Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. *Med Sci Sports Exerc.* 2007;39:1423–1434.
36. Crespo CJ, Smit E, Anderson RE, et al. Race/ethnicity, social class and their relation to physical inactivity during leisure time: results from the Third National Health and Nutrition Examination Survey, 1988–1994. *Am J Prev Med.* 2000;18:46–53.
37. Anderson RT, King A, Stewart AL, Camacho F, Rejeski WJ. Physical activity counseling in primary care and patient well-being: do patients benefit? *Ann Behav Med.* 2005;30:146–154.
38. Pinto BM, Friedman R, Marcus BH, et al. Effects of a computer-based telephone counseling system on physical activity. *Am J Prev Med.* 2002;23:113–120.
39. Tse WS, Bond AJ. The impact of depression on social skills: a review. *J Nerv Ment Dis.* 2004;192:260–268.
40. Eyler AA, Matson-Koffman D, Young DR, et al. Quantitative study of correlates of physical activity in women from diverse racial/ethnic groups: The Women's Cardiovascular Health Network Project summary and conclusions. *Am J Prev Med.* 2003;25:93–103.
41. Eyler AA, Brownson RC, Donatelle RJ, King AC, Brown D, Sallis JF. Physical activity social support and middle- and older-aged minority women: results from a US survey. *Soc Sci Med.* 1999;49:781–789.
42. Wilbur J, Miller MA, Chandler P, McDevitt J. Determinants of physical activity and adherence to a 24-week home-based walking program in African American and Caucasian women. *Res Nurs Health.* 2003;26:213–224.

Behavioral Assessments in Russian Addiction Treatment Inpatients: A Comparison of Audio Computer-Assisted Self-Interviewing and Interviewer-Administered Questionnaires

Erika M. Edwards,[1] Debbie M. Cheng,[2] Suzette Levenson,[1] Carly Bridden,[3] Seville Meli,[3] Valentina Y. Egorova,[4] Evgeny M. Krupitsky,[4] and Jeffrey H. Samet[3,5]

[1]Boston University School of Public Health, Data Coordinating Center, Boston, Massachusetts, USA; [2]Boston University School of Public Health, Department of Biostatistics, Boston, Massachusetts, USA; [3]Boston University School of Medicine/Boston Medical Center, Department of Medicine, Section of General Internal Medicine, Clinical Addiction Research and Education (CARE) Unit, Boston, Massachusetts, USA; [4]St. Petersburg State Pavlov Medical University, St. Petersburg, Russian Federation; [5]Boston University School of Public Health, Department of Social and Behavioral Sciences, Boston, Massachusetts, USA

Purpose: Assess agreement between reported sex and drug use behaviors from audio computer-assisted self-interviewing (ACASI) and interviewer-administered questionnaire (IAQ). **Method:** Participants (N = 180) enrolled in an HIV intervention trial in Russia completed ACASI and IAQ on the same day. Agreement between responses was evaluated. **Results:** Of the 13 sex behavior questions, 10 items had excellent agreement (kappas/ICC 0.80–0.95) and 3 items had moderate agreement (kappas/ICC 0.59–0.75). The 3 drug behavior questions had excellent agreement (kappas/ICC 0.94–0.97). Among HIV-specific questions asked of HIV-positive participants (n = 21) only, 2 items had excellent agreement (kappas 1.0) and 3 items had moderate agreement (kappas 0.40–0.71). **Conclusions:** Assessment of drug and sex risk behaviors by ACASI and IAQ had generally strong agreement for the majority of items. The lack of discrepancy may result from these Russian subjects' perception that computers do not ensure privacy. Another potential explanatory factor is that both interviews were delivered on the same day. These data raise questions as to whether use of ACASI is uniformly beneficial in all settings, and what influence cultural factors have on its utility. **Key words:** *HIV infections, interview methods, prevention & control, Russia*

Obtaining information on self-reported risk behaviors related to HIV transmission has been both a critical and a challenging factor throughout the HIV epidemic. In HIV-related studies, patients may need to disclose sensitive, embarrassing, or even illegal information about sex and drug-related activities. Some individuals may deny engaging in what are perceived to be undesirable behaviors, resulting in social desirability bias.[1] This bias could over- or underestimate the effectiveness of intervention programs. Therefore, truthful reporting of risk behaviors is crucial for unbiased assessments.

Computer-assisted self-interview (CASI) and audio computer-assisted self-interviewing (ACASI) allow patients to complete questionnaires on their own via a computer. In an ACASI system, patients read the questions on a screen and listen to them through headsets, reducing potential literacy barriers.[2] Traditional methods of collecting sensitive information, such as interviewer-administered questionnaires (IAQ), require direct or indirect involvement by research staff. Research participants are thought to overreport desirable behaviors or underreport undesirable behaviors when interviews are done in person.[2]

Newer methods, such as CASI and ACASI, may increase privacy and reduce reporting bias compared to IAQ,[3] although in the United States the evidence is mixed.[2] For example, among sexually transmitted infection clinic patients who com-

Address for correspondence: Erika M. Edwards, MPH, Data Coordinating Center, Boston University School of Public Health, 715 Albany Street, Crosstown Center, 3rd Floor, Boston, MA 02118 USA. Email: eedwards@bu.edu

HIV Clin Trials 2008;9(4):247–253

www.thomasland.com

doi: 10.1310/hct0904-247

pleted both ACASI and IAQ, participants reported higher rates of sex risk behaviors to ACASI but there was no difference in rates of drug risk behaviors[4]; similar results were found among sexually transmitted infection clinic patients who were randomized to ACASI or IAQ.[5] Studies of drug users generally have shown more reporting of sensitive behaviors to ACASI than IAQ.[6–8]

Use of CASI and ACASI systems has been increasing internationally in diverse populations such as adolescents in Vietnam,[9] Kenya,[10] and India[11]; alcohol and drug users in Brazil[12]; women in Zimbabwe[13] and Kenya[14]; urban market workers in China[15]; and community volunteers in China, India, Peru, Russia, and Zimbabwe.[16] All of these studies compared reporting by ACASI to IAQ. Four found higher prevalence of sensitive behaviors among those using ACASI. Adolescents in Vietnam who were randomized to ACASI were more likely to report risky sexual practices than those who were randomized to IAQ.[9] Adolescent girls in Kenya who were randomized to ACASI reported significantly different sexual practices than girls who were randomized to IAQ.[10] Drug users in Brazil who were randomized to ACASI were more likely to report multiple drug use and risky sexual behaviors compared to those who were randomized to IAQ.[11] Finally, breastfeeding women in Kenya who were assigned to complete both ACASI and IAQ in a randomized crossover design were more likely to report sensitive behaviors such as less time breastfeeding or earlier introduction of complementary foods to ACASI.[14]

The other three studies found mixed results. In India, adolescent girls were more likely to underreport sensitive sexual behaviors in ACASI than in IAQ. Adolescent boys' responses depended upon the type of sexual behavior; some were overreported in ACASI compared to IAQ, some underreported, and some the same.[11] In Zimbabwe, 76% of women had no differences between ACASI and IAQ responses.[13] Urban market workers in China were more likely to report engaging in lifetime sexual intercourse during IAQ than ACASI, but there were no differences in the number of lifetime sexual partners or in responses to sexually transmitted disease (STD)-related questions.[15] Finally, in a multi-country comparison of computer-assisted personal interviewing and ACASI, concordance varied by country, although most participants' responses did not differ by mode.[16] In all of these studies, participants completed both ACASI and IAQ but were randomized to which they completed first.

Russia has the fastest growing HIV epidemic in Europe, with an estimated 940,000 infected since the mid 1990s.[17] HIV transmission was initially predominantly through injection drug use[18] but is spreading to the general population via sexual transmission.[17,19] Given the importance of the HIV epidemic in Russia, further research about disease transmission in this country will require understanding the advantages and disadvantages of research methodologies assessing risky behaviors.

The objective of this study was to assess the agreement between responses obtained from ACASI and those obtained from face-to-face interviews in Russian participants who completed both interview modes. Specifically, our goal was to evaluate whether participants would be more likely to overreport desired behaviors and underreport undesirable behaviors in face-to-face interviews compared to ACASI. Given the awareness of past collection of personal information on individuals with detrimental consequences or incomplete privacy assurances in some Eastern European countries, it was hypothesized that the use of recording devices would not yield more revealing information in the Russian setting. We compared responses to sexual and drug use behavior questionnaires collected as part of the Russian PREVENT Study (Partnership to Reduce the Epidemic Via Engagement in Narcology Treatment).[20]

METHOD

Study Design and Participants

The Russian PREVENT study was a randomized controlled trial of men and women with alcohol and/or drug dependence recruited from two inpatient substance abuse treatment facilities (narcology hospitals) near St. Petersburg, Russia. Trained physician research associates approached patients after initial detoxification to assess eligibility, offer participation, and conduct assessments. Criteria for participant eligibility included the following: age 18 years and older, a primary diagnosis of alcohol or drug dependence, no alcohol and other abused substances for at least 48 hours, reported unprotected anal or vaginal sex in the past 6 months, willing to undergo HIV testing

as per standard narcology hospital counseling and testing protocol or previously diagnosed as HIV infected, able to provide reliable contact information including a home telephone number, and an address within 150 kilometers of St. Petersburg. Patients not fluent in Russian or with cognitive impairment based on the research associates' clinical judgment were excluded from the study. All participants provided written informed consent prior to enrollment in the study. The Institutional Review Boards of Boston Medical Center and St. Petersburg Pavlov State Medical University approved this study. PREVENT had 181 enrolled participants, but 1 person was missing ACASI responses and therefore was dropped from these analyses. The current analyses focus on data collected at the baseline assessment.

Participant Assessment

Participant recruitment and follow-up occurred from October 2004 through December 2005. Baseline assessments measured risk behaviors by IAQ and an ACASI system. Interviewers were blinded to intervention group. All interviews were conducted in Russian, including those done by the ACASI system. Participants were compensated the equivalent of US $5 for the baseline assessment.

Instrument Design and Data Collection

Programmers at the Boston University School of Public Health Data Coordinating Center developed the ACASI system in Microsoft Access. Paper forms were created in English, translated into Russian by the Russian investigators, and back-translated by the Data Coordinating Center. The audio track was recorded with a male and a female voice, and the participants were allowed to choose which they preferred. Participants completed both the IAQ and the ACASI on the same day. There was not a protocol for a consistent pattern to whether the ACASI or the IAQ occurred first.

Interventionists (two psychiatrists and a psychologist trained in HIV and addiction) were trained to deliver the intervention by US collaborators using a standard curriculum. The lead interventionist underwent an initial training in English in the United States. A subsequent 3-day training in St. Petersburg with simultaneous translation allowed multiple role-playing sessions to be observed and critiqued by a behavioral psychologist on the US study team.

Questions about sex and drug risk behaviors came from the RESPECT study (has primary sex partner, has secondary sex partner, mean number of times unprotected vaginal sex with primary partner past 3 months, mean number of times unprotected vaginal sex with secondary partner past 3 months[21]) and the Risk Assessment Battery ([RAB]; heterosexual, multiple sex partners past 6 months, 4+ opposite sex partners past 6 months, same sex partner ever, buy sex, sell sex, used condoms past 6 months, unsafe drug use past 6 months, shared needles past 6 months, mean Drug Risk Assessment score[22]). Questions on HIV disclosure, asked only of HIV-infected subjects, did not come from a specific instrument.

Statistical Analysis

Descriptive statistics such as means, medians, standard deviations, and proportions were used to describe the study population. Agreement between responses obtained from IAQ and ACASI was assessed using the kappa statistic for dichotomous variables and the intraclass correlation coefficient (ICC) for continuous variables. We used the following guidelines to interpret values of kappa and ICC: <0.4, poor agreement; 0.4–0.75, moderate agreement; >0.75, excellent agreement.[23] In addition, we compared participant responses from the two interview methods using McNemar's test for paired dichotomous variables and the Wilcoxon signed rank test for paired continuous variables. Exact p values were used for McNemar's test when the number of discordant pairs was <20. Additional exploratory analyses were conducted stratifying by drug diagnosis (yes vs. no) and HIV status (positive vs. negative). All analyses included only data collected at the baseline visit.

RESULTS

Study participants, described in **Table 1**, included the following: 75% male, 94% high school graduates, 33% married, and 15% HIV infected. Responses to sex behavior questions by ACASI and IAQ are presented in **Table 2**. Of the 13 sex behavior questions, 10 out of 13 (77%) items (heterosexual, primary or secondary sex partners, multiple sex partners, 4+ opposite sex partners, same sex part-

Table 1. Baseline characteristics of Russian narcology patients from the PREVENT cohort ($N = 180$)

Characteristic	
Median age, years (IQR)	30 (26–40)
Male	135 (75)
Employed full time	89 (49)
High school graduate	169 (94)
Married	60 (33)
Diagnosis	
Alcohol	107 (59)
Heroin	58 (32)
Alcohol and heroin	15 (8)
HIV infected	27 (15)

Note: Data are number (%) unless otherwise specified. IQR = interquartile range.

ner, sex trade, any sexually transmitted diseases, any condom use) had excellent agreement (kappas 0.80–0.95). Three out of 13 (23%) items (prior HIV testing and number of unprotected sex episodes with primary and secondary partners) had moderate agreement (kappas/ICC 0.59–0.75).

The question "Prior to this hospitalization, have you ever been tested for HIV?" had a kappa statistic of 0.59. Of the 17 subjects with discordant responses on the ACASI and IAQ for this question, 13 people (76%) responded "yes" in the ACASI and "no" in the IAQ and 4 people (24%) responded "no" in the ACASI and "yes" in the IAQ, a statistically significant finding (McNemar $p < 0.03$). The question "Do you have a primary sex partner?" had a kappa of 0.86, suggesting excellent agree-

Table 2. Baseline sex behaviors reported by audio computer-assisted self-interviewing (ACASI) and interviewer-administered questionnaire (IAQ) interview modes from 180 Russian narcology patients

	IAQ	ACASI	Agreement[a]	McNemar *p* value	Wilcoxon signed rank test *p* value
Heterosexual	96%	97%	0.95	0.56	
Primary sex partner	79%	76%	0.86	0.02	
Secondary sex partner	69%	67%	0.83	0.17	
Multiple sex partners, past 6 months	69%	69%	0.91	0.71	
4+ opposite sex partners, past 6 months	28%	26%	0.82	0.37	
Same sex partner, ever	7%	9%	0.89	0.08	
Buy sex	19%	19%	0.80	0.76	
Sell sex	9%	12%	0.85	0.03	
STDs, ever	43%	43%	0.90	0.74	
Used condoms, past 6 months	60%	56%	0.80	0.06	
Prior to this hospitalization, ever tested for HIV	89%	84%	0.59	0.03	
Number of times unprotected vaginal sex with primary partner, past 3 months ($n = 122$[b])			0.75		0.95
Mean (*SD*)	26.2 (23.7)	26.5 (23.6)			
Median (IQR)	17.0 (8–39)	20.0 (8–40)			
Number of times unprotected vaginal sex with secondary partner, past 3 months ($n = 103$[b])			0.74		0.85
Mean (*SD*)	9.1 (12.2)	10.0 (14.7)			
Median (IQR)	5.0 (1–10)	5.0 (2–10)			

Note: STDs = sexually transmitted diseases; IQR = interquartile range.

[a]Agreement for categorical variables measured with the kappa statistic and for continuous variables with the intraclass correlation coefficient.

[b]Number of matched pairs with data for both questions.

ment. Among the nine discordant responses for this question, eight people (89%) responded "yes" in the ACASI and "no" in the IAQ and one person (1%) responded "no" in the ACASI and "yes" in the IAQ. McNemar's test was significant ($p < 0.02$), indicating that in instances of discordance participants were more likely to respond affirmatively with the ACASI.

Responses to drug use behavior questions by ACASI and IAQ are presented in **Table 3**. All three of the responses, which pertained to unsafe drug use, needle sharing, and the Drug Risk Assessment score, had excellent agreement (kappas/ICC 0.94–0.97).

The 21 HIV-infected participants were asked about disclosure of their serostatus; the 159 HIV-negative participants were not asked these questions (**Table 4**). Two (60%) questions had excellent agreement and three (40%) questions had moderate agreement. Twenty participants (95%) had concordant responses to the ACASI and IAQ when asked if they had ever told anyone if they were HIV infected and if they had told family members of their serostatus. The question about disclosing to friends had a kappa statistic of 0.70. The question about disclosing to any sexual partners in the past 6 months had a kappa statistic of 0.71. For each of the three discordant responses to these questions, participants responded "no" in the ACASI and "yes" in the IAQ. Finally, the question about disclosing to all sexual partners in the past 6 months had a kappa statistic of 0.40 and a significant McNemar's test p value ($p < 0.01$). Among the seven discordant responses, each responded "no" in the ACASI and "yes" in the IAQ.

In secondary analyses stratified by drug diagnosis and HIV status, agreement between the two interview modes appeared similar across subgroups (data not shown).

DISCUSSION

ACASI systems are used in research studies to improve truth telling concerning risky and potentially stigmatized behaviors.[2] ACASI has been used increasingly in recent years in HIV-related research studies in lieu of IAQ to improve privacy and reduce participants' inclination to give socially desirable responses. Seven published studies performed outside the United States compared responses to ACASI versus IAQ. Three found increased reporting of socially undesirable or stigmatized behaviors via ACASI.[9,10,12] These studies performed in Brazil, Vietnam, and Kenya compared the ACASI information to interview by randomizing participants to one method or the other. Four found mixed results.[11,13,14,15] In all of these studies, participants completed both ACASI and IAQ but were randomized to which they completed first.

For the most part, no strong reporting bias in responses to sexual and drug behavior questions was observed based on assessment methodology in this study of Russians with alcohol and/or drug dependence. There was excellent agreement between responses from ACASI and IAQ for the majority of responses.

A moderate level of agreement was observed on some questions related to HIV testing overall and on serostatus disclosure among HIV-infected participants. All participants were asked, "Prior

Table 3. Baseline drug use behaviors reported by audio computer-assisted self-interviewing (ACASI) and interviewer-administered questionnaire (IAQ) interview modes from 180 Russian narcology patients

	IAQ	ACASI	Agreement[a]	McNemar *p* value	Wilcoxon signed rank test *p* value
Unsafe drug use, past 6 months ($n = 178$[b])	30%	31%	0.97	0.16	
Shared needles, past 6 months ($n = 80$[b])	67%	70%	0.94	0.32	
Drug Risk Assessment Score ($n = 80$[b])			0.94		0.70
Mean (*SD*)	11.0 (7.2)	11.5 (6.9)			
Median (IQR)	11 (4–17)	12 (5.5–17)			

[a]Agreement for categorical variables measured with the kappa statistic and for continuous variables with the intraclass correlation coefficient.

[b]Number of matched pairs with data for both questions.

Table 4. Baseline HIV disclosure behaviors reported by audio computer-assisted self-interviewing (ACASI) and interviewer-administered questionnaire (IAQ) interview modes.

	IAQ	ACASI	Agreement[a]	McNemar test *p* value
Told anyone you are HIV infected	95%	95%	1.0	1.0
Disclosed HIV status to family	95%	95%	1.0	1.0
Disclosed HIV status to friends	62%	57%	0.70	0.56
Disclosed HIV status to any sex partners in past 6 months	67%	52%	0.71	0.08
Disclosed HIV status to all sex partners in past 6 months	43%	33%	0.40	<0.01

Note: Questions asked only of participants who tested positive for HIV ($n = 21$).
[a]Agreement measured with the kappa statistic.

to this hospitalization, have you ever been tested for HIV?" Participants with discordant responses were more likely to report "yes" to ACASI and "no" to the interviewer (McNemar $p < 0.03$). The HIV epidemic is of more recent onset in Russia compared to other countries. This area of discordance between interview approaches may be due to perceived stigma of being HIV infected and increased willingness to report a private experience such as HIV testing to ACASI. Additionally, the 21 HIV-infected participants were asked questions about disclosing their serostatus. When asked about disclosing to any or all sexual partners in the past 6 months, discordant participants were more likely to report "yes" to the interviewer and "no" to the ACASI. This discordance may be due to overreporting desirable behaviors to interviewers or to increased willingness to report less desirable behaviors to ACASI.

Overall, however, we observed strong agreement in the majority of items between the two interview modes. Two factors may explain this observation. The first is the study design, where participants completed both interview modes on the same day rather than being randomly assigned to a single interview mode. The participants, who already answered questions about their substance use and sexual behaviors to one mode of interview (IAQ or ACASI), may have provided consistent responses because they believed their answers would be compared.

A second factor is the potential perception of the study participants that computers do not ensure privacy. In Russia, databases with personal data on individual bank accounts, mobile phone numbers, passport information, and more are available for sale on the street. Many Russians do not believe that personal data reported to a computer and stored in the electronic form are more confidential and safer compared to paper form, especially because electronic information can be easily copied and distributed. In another study comparing ACASI and IAQ, only 23% of Russians responded "yes, absolutely" when asked whether the computer ensures sufficient privacy compared to 40% of participants in China, India, and Peru.[16] This finding may be indicative of a larger mistrust of computers in Russia. Assessing study participants' preference of ACASI or IAQ would provide insight about use of this methodology in Russia.

In contrast to previous studies in the United States and other settings comparing ACASI and IAQ, this study observed generally strong agreement between responses to the two interview modes. The ACASI does not appear to substantially reduce reporting bias, compared to an IAQ, among inpatients in a substance abuse treatment facility in Russia. The advantages of advanced research techniques such as ACASI in one cultural setting may not be generalizable to other cultural settings. Based on this experience, the need to use the ACASI methodology in Russia for assessing risky behaviors is not compelling. Further studies utilizing a randomized controlled study design to compare responses from ACASI and IAQ in Russia would need to demonstrate a clear benefit of the ACASI methodology in order to make the case for these "advanced" methods to be utilized in Russia.

ACKNOWLEDGMENTS

Grant support for this study came from the National Institute on Alcohol Abuse and Alcoholism

(NIAAA), NIH R21-AA014821 (HIV Prevention Partnership in Russian Alcohol Treatment), and K24-AA015674 (Impact of Alcohol Use on HIV Infection—In US and Russia).

We would like to thank the interventionists, research associates, and participants at the Russian narcology hospitals for their contributions to the study, and Denis Rybin of the Data Coordinating Center for providing back translation of the study forms.

REFERENCES

1. Rosenthal R, Persigner GW, Fode KL. Experimental bias, anxiety, and social desirability. *Percept Mot Skills.* 1962; 15:73–74.
2. Newman JC, Des Jarlais DC, Turner CF, Gribble J, Cooley P, Paone D. The differential effects of face-to-face and computer interview modes. *Am J Public Health.* 2002;92(2):294–297.
3. O'Reilly J, Hubbard M, Lessler P, et al. Audio and video computer-assisted self-interviewing: preliminary tests of new technologies for data collection. *J Off Stat.* 1994;10: 197–214.
4. Ghanem KG, Hutton EH, Zenilman JM, Zimba R, Erbelding EJ. Audio computer assisted self interview and face to face interview modes in assessing response bias among STD clinic patients. *Sex Transm Infect.* 2005;81:421–425.
5. Rogers SM, Willis G, Al-Tayyib A, et al. Audio computer assisted interviewing to measure HIV risk behaviours in a clinic population. *Sex Trans Infect.* 2005;81:501–507.
6. Des Jarlais DC, Paone D, Milliken J, et al. Audio-computer interviewing to measure risk behaviour for HIV among injecting drug users: a quasi-randomised trial. *Lancet.* 1999; 35:1657–1662.
7. Macalino GE, Celentano DD, latkin C, Strathdee SA, Vlahov D. Risk behaviors by audio computer-assisted self-interviews among HIV-seropositive and HIV-seronegative injection drug users. *AIDS Educ Prev.* 2002;14:367–378.
8. Perlis TE, Des Jarlais DC, Friedman SR, Arasteh K, Turner CF. Audio-computerized self-interviewing versus face-to-face interviewing for research data collection at drug abuse treatment programs. *Addiction.* 2004;99:885–896.
9. Le LC, Blum RW, Magnani R, Hewett PC, Do HM. A pilot of audio computer-assisted self-interview for youth reproductive health research in Vietnam. *J Adolesc Health.* 2006;38:740–747.
10. Hewett PC, Mensch BS, Erulkar AS. Consistency in the reporting of sexual behaviour by adolescent girls in Kenya: a comparison of interviewing methods. *Sex Trans Infect.* 2004;80(Suppl II):ii43–ii48.
11. Jaya, Hindin MJ, Ahmed S. Differences in young people's reports of sexual behaviors according to interview methodology: a randomized trial in India. *Am J Public Health.* 2008;98:169–174.
12. Simoes AA, Bastos FI, Moreira RI, Lynch KG, Metzger DS. A randomized trial of audio computer and in-person interview to assess HIV risk among alcohol and drug users in Rio De Janeiro, Brazil. *J Subst Abuse Treat.* 2006;30:237–243.
13. van de Wijert J, Padian N, Shiboski S, Turner C. Is audio computer-assisted self-interviewing a feasible method of surveying in Zimbabwe? *Int J Epidemiol.* 2000;29:885–890.
14. Waruru AK, Nduati R, Tylleskär T. Audio computer-assisted self-interviewing (ACASI) may avert socially desirable responses about infant feeding in the context of HIV. *BMC Med Inform Decis Mak.* 2005;5:24.
15. Li L, Wu Z, Rotheram-Borus MJ, et al. Measuring sexual risk using audio computer-assisted self-interviewing (ACASI) versus computer-assisted personal interview (CAPI) in China. *Int J Sexual Health.* 2007;19(1):25–30.
16. NIMH Collaborative HIV/STD Prevention Trial Group. The feasibility of audio computer-assisted self-interviewing in international settings. *AIDS.* 2007;21(Suppl 2):S49–S58.
17. UNAIDS. Russian Federation Report. Available at: http://www.unaids.org/en/Regions_Countries/Countries/russian_federation.asp
18. Dehne KL, Pokrovskiy V, Kobyshcha Y, Schwartlander B. Update on the epidemics of HIV and other sexually transmitted infections in the newly independent states of the former Soviet Union. *AIDS.* 2000;14(Suppl 3):S75–S84.
19. Hamers FF, Downs AM. HIV in central and eastern Europe. *Lancet.* 2003;361:1035–1044.
20. Samet JH, Krupitsky EM, Cheng DM, et al. Mitigating risky sexual behavior among Russian narcology hospital patients: the PREVENT (Partnership to Reduce the Epidemic via Engagement in Narcology Treatment) randomized control trial. *Addiction.* Epub July 10, 2008.
21. Kamb ML, Fishbein M, Douglas JM Jr, et al. Efficacy of risk reduction counseling to prevent human immunodeficiency virus and sexually transmitted diseases: a randomized controlled trial. Project RESPECT Study Group. *JAMA.* 1998;280(13):1161–1167.
22. Navaline HA, Snider EC, Petro CJ, et al. Preparations for AIDS vaccine trials. An automated version of the Risk Assessment Battery (RAB): enhancing the assessment of risk behaviors. *AIDS Res Hum Retroviruses.* 1994;10(Suppl 2):281–283.
23. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159–174.

Relations of Thyroid Function to Body Weight

Cross-sectional and Longitudinal Observations in a Community-Based Sample

Caroline S. Fox, MD, MPH; Michael J. Pencina, PhD; Ralph B. D'Agostino, PhD; Joanne M. Murabito, MD; Ellen W. Seely, MD; Elizabeth N. Pearce, MD; Ramachandran S. Vasan, MD

Background: Overt hypothyroidism and hyperthyroidism may be associated with weight gain and loss. We assessed whether variations in thyroid function within the reference (physiologic) range are associated with body weight.

Methods: Framingham Offspring Study participants (n = 2407) who attended 2 consecutive routine examinations, were not receiving thyroid hormone therapy, and had baseline serum thyrotropin (TSH) concentrations of 0.5 to 5.0 mIU/L and follow-up concentrations of 0.5 to 10.0 mIU/L were included in this study. Baseline TSH concentrations were related to body weight and body weight change during 3.5 years of follow-up.

Results: At baseline, adjusted mean weight increased progressively from 64.5 to 70.2 kg in the lowest to highest TSH concentration quartiles in women ($P < .001$ for trend), and from 82.8 (lowest quartile) to 85.6 kg (highest quartile) in men ($P = .007$ for trend). During 3.5 years of follow-up, mean (SD) body weight increased by 1.5 (5.6) kg in women and 1.0 (5.0) kg in men. Baseline TSH concentrations were not associated with weight change during follow-up. However, an increase in TSH concentration at follow-up was positively associated with weight gain in women (0.5-2.3 kg across increasing quartiles of TSH concentration change; $P < .001$ for trend) and men (0.4-1.3 kg across quartiles of TSH concentration change; $P = .007$ for trend).

Conclusions: Thyroid function (as assessed by serum TSH concentration) within the reference range is associated with body weight in both sexes. Our findings raise the possibility that modest increases in serum TSH concentrations within the reference range may be associated with weight gain.

Arch Intern Med. 2008;168(6):587-592

Author Affiliations: National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, Massachusetts (Drs Fox, Pencina, D'Agostino, Murabito, and Vasan); the National Heart, Lung, and Blood Institute, Bethesda, Maryland (Dr Fox); Department of Endocrinology, Hypertension, and Diabetes, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts (Drs Fox and Seely); Department of Mathematics, Boston University (Drs Pencina and D'Agostino), and Section of General Internal Medicine (Dr Murabito), Department of Endocrinology (Dr Pearce), Division of Preventive Medicine (Dr Vasan), and Cardiology Section (Dr Vasan), Boston University School of Medicine, Boston, Massachusetts.

OBESITY IS ASSOCIATED WITH an increased risk for diabetes,[1-3] vascular disease,[4] all-cause mortality,[5] and cancer.[6] The prevalence of obesity has increased substantially in the United States,[7,8] suggesting a need to understand risk factors for weight gain. There are multiple known predictors of obesity and weight gain, including a low level of physical activity,[9] increased caloric intake,[10] parity,[11] smoking cessation,[11] inflammation,[12] depression,[13] and genetic factors.[14] In addition, metabolic factors are associated with body weight, including biomarkers of adiposity

For editorial comment see page 568

(adipocytokines)[15] and a lower resting metabolic rate.[16] In this context, it is noteworthy that thyroid dysfunction is well recognized as a cause of weight change. Weight loss is a frequent manifestation of hyperthyroidism, and hyperthyroid patients who are treated adequately gain nearly 4 kg/y.[17] Conversely, weight gain is a common complaint in patients with hypothyroidism, and treatment with thyroid hormone is associated with modest weight loss.[18]

However, trials of subclinical hypothyroidism have not resulted in significant weight loss,[18,19] raising the question of whether body weight varies with variation in thyroid function within the reference (physiologic) range. Some studies have related thyroid function and body weight in small, selected samples of obese individuals[20,21] or those with thyroid disease.[22] Some,[20,21] but not all,[22] have reported an association between thyroid function and body weight. Although variation in thyroid function within the reference range has been related to weight change in population-based samples, prior studies were limited by the use of self-reported weight gain[23] and the lack of an adequate control for baseline weight[24] (a known risk factor for weight gain[25]) in multivariable analyses. Thus, it is not known whether variation in thyroid function within the reference range is associated with body weight, or whether subtle changes in thyroid function are associated with changes in weight over time. In this study, we determined the relations be-

tween thyroid function, as assessed by measurement of serum thyrotropin (TSH) concentrations, and body weight and weight change in a large community-based cohort. We hypothesized, based on knowledge of the fundamental role of thyroid hormone in regulating metabolism, that higher levels of serum TSH within the reference range may be associated with lesser degrees of thyroid hormone activity, and consequently with greater cross-sectional body weight. Furthermore, we postulated that incremental changes in serum TSH concentrations over time may be associated with longitudinal weight gain.

METHODS

STUDY SAMPLE

Participants for this study were drawn from the Framingham Heart Study. Beginning in 1948, 5209 men and women aged 28 to 62 years were enrolled.[19] Offspring of the original cohort and their spouses (n=5124) were enrolled in the Offspring Study cohort starting in 1971. The selection criteria and design of the Offspring Study have been described previously.[26,27] The study sample for the present study consisted of 3583 Offspring Study cohort participants who attended the third (*baseline*; 1983-1987) and fourth (*follow-up*; 1987-1991) examinations. Of these, 2407 were available for the final analyses after the following exclusions: use of thyroid hormone therapy at baseline (n=121) and at follow-up (n=40); missing covariates (n=92); missing serum TSH values (n=736); serum TSH concentration of less than 0.5 mIU/L (n=52) or higher than 5.0 mIU/L at baseline (n=108); or serum TSH concentrations of less than 0.5 mIU/L (n=20) or higher than 10.0 mIU/L at follow-up (n=7). Eligibility thresholds based on serum TSH values were slightly more liberal at the follow-up examination to ensure inclusion of participants who had moderate changes in serum TSH concentrations over time. The participants with available serum TSH values were similar to those with missing values with respect to body weight and body mass index. More women than men had missing TSH values (24.5% vs 19.7%; $P<.001$). Men with missing serum TSH values were slightly older than men with serum TSH values (51.0 vs 48.0 years; $P<.001$); there were no differences among women.

SERUM TSH ASSAY

Concentrations of TSH were measured on fasting morning samples using a chemoluminescence assay (London Diagnostics, Eden Prairie, Minnesota) with a lower limit of detection of 0.01 mIU/L. The interassay coefficients of variation were 5% (at 1 mIU/L) and 11% (at 0.04 mIU/L) as previously described.[28] Serum-free thyroxine levels were not measured.

RISK FACTOR ASSESSMENT

At each examination, weight was measured to the nearest pound with the participant wearing only a gown without shoes or slippers and standing in the middle of the scale (Detecto scale; Worchester Scale Co Inc, Webb City, Massachusetts) with the weight equally distributed on both feet. The scale is calibrated with a 22.5-kg weight monthly and professionally calibrated annually. Smoking status was defined by tobacco use in the year preceding the examination. For longitudinal analyses, the following 3-level smoking variable was defined: never, current, or quit in the interim. Women were considered to be postmenopausal if their menstrual periods had stopped for at least 1 year; a separate 3-level variable was constructed for baseline, interim, or no menopause for longitudinal analyses.

STATISTICAL ANALYSIS

We decided a priori to perform all analyses separately in men and women because of the well-known differences in thyroid disease prevalence in women compared with men.[29] Our TSH assay varied over time; therefore, statistical calibration to the Third National Health and Nutrition Examination Survey (NHANES III) data[29] was carried out using log-normal transformations within sex-specific age groups. Briefly, we divided women and men in our sample into 5-year age groups (30-34, 35-39, 40-44, etc) and standardized their log-transformed values within each group. Then we imposed NHANES III–based means, medians, and standard deviations corresponding to the log-transformed values. This was achieved by first multiplying the log-transformed standardized values by the standard deviation of the log-transformed log-normal variable calculated as

$$\sqrt{\log[(\sigma/\mu-\lambda)^2+1]},$$

where μ and σ denote the NHANES III–based log-normal means and standard deviations and λ is the shift variable calculated using the NHANES III–based median, minimum, and maximum:

$$(\text{Minimum}\times\text{Maximum}-\text{Median Squared})/(\text{Minimum}+\text{Maximum}-2\times\text{Median})$$

and adding the corresponding mean calculated as

$$\log[(\mu-\lambda)^2/\sqrt{\sigma^2+(\mu-\lambda)^2}].$$

The resulting number was then exponentiated and a shift parameter λ was added to obtain the final value. This method preserved not only the age-group and sex-specific means and standard deviations but also the medians.

We analyzed TSH concentrations as a continuous variable and as quartiles. When used as a continuous variable, TSH concentration was natural logarithmically transformed because of its skewed distribution; results are presented as weight change per 1-unit increase in log TSH concentration. To formally test for differences between women and men, we pooled the sexes and fit a sex interaction term in the multivariable models on sex-standardized weight. For quartile-based analyses, we examined sex-specific quartiles of serum TSH concentrations in relation to weight and weight change. Continuous and quartile-based analyses of TSH concentrations consisted of the following 3 models:

1. Cross-sectional association of serum TSH concentrations and body weight at the baseline examination using multivariable linear regression models, adjusting for age, smoking, and menopausal status.
2. Relations of baseline serum TSH concentrations to the longitudinal change in body weight at follow-up, adjusting for age, 3-level smoking status, 3-level menopause status, and baseline weight using multivariable linear regression.
3. Sex-specific association between change in serum TSH concentration and change in body weight at follow-up, adjusting for age, 3-level smoking status, 3-level menopause status, and baseline weight using multivariable linear regression.

Secondary analyses were performed to determine whether participants in the highest decile of weight change were more likely to have elevated serum TSH values at follow-up compared with the rest of the sample. For these analyses, we relaxed our eligibility criteria to permit inclusion of participants regardless of their serum TSH value to increase the clinical relevance; however, participants reporting use of thyroid hormone therapy were not included because we were unable to assess recent medication changes. Two sex-specific analyses were performed, adjusting only for age: determination of the log-transformed mean serum TSH

concentration at follow-up in participants in the upper decile of weight change at follow-up compared with the remainder of the sample; and determination of the odds ratio of having a serum TSH concentration higher than 5.0 mIU/L among participants in the upper decile of weight change at follow-up compared with the participants in the lower 9 deciles.

For all analyses, a 2-tailed $P < .05$ was considered significant. We used SAS, version 9.1 statistical software to perform all computations.[30]

RESULTS

Overall, data for 1117 women and 1290 men were available for analysis. The average interval between the baseline and follow-up examinations was 3.5 years. The women weighed an average of 66.6 kg at the baseline examination and 68.0 kg at follow-up (**Table**). The men weighed an average of 84.1 kg at baseline and 85.1 kg at follow-up. The prevalence of obesity at baseline was 14.3%. The mean baseline serum TSH concentration was 1.91 mIU/L in women and 1.70 mIU/L in men.

CROSS-SECTIONAL RELATIONS OF BASELINE SERUM TSH CONCENTRATIONS AND BODY WEIGHT

In analyses modeling TSH concentration as a continuous variable, an increase of 1 U of log TSH concentration was associated with a 4.2-kg greater weight in women ($P < .001$) and 1.9-kg greater weight in men ($P = .01$); the sex interaction term was of borderline statistical significance ($P = .07$). In quartile-based analyses, there was a similar strong and graded positive relation between increasing quartiles of serum TSH concentrations and higher body weight at baseline. Women with the lowest quartile of TSH concentration had a mean weight of 64.5 (95% confidence interval [CI], 62.9-66.1) kg, whereas women in the highest quartile of TSH concentration had a mean weight of 70.2 (95% CI, 68.5-71.8) kg (**Figure 1**; $P < .001$ for trend). Men in the lowest quartile of TSH concentration had a mean weight of 82.8 (95% CI, 81.4-84.2) kg, whereas those in the highest quartile had a mean weight of 85.6 (95% CI, 84.2-86.9) kg ($P = .007$ for trend).

BASELINE SERUM TSH CONCENTRATION AND WEIGHT CHANGE AT FOLLOW-UP

There was no relation between log TSH concentration (modeled as a continuous variable) and weight change during follow-up in women ($P = .25$) or men ($P = .90$), and no sex interaction was observed ($P = .33$ for the interaction term). In the quartile-based analysis, baseline serum TSH concentration was not related to weight change during the 3.5 years of follow-up in either sex (**Figure 2**).

CHANGE IN SERUM TSH CONCENTRATION AND WEIGHT CHANGE AT FOLLOW-UP

At follow-up, the changes in serum TSH concentrations (ΔTSH) ranged from –2.45 to 7.49 mIU/L in women and –3.02 to 6.79 mIU/L in men. In women, weight increased by 2.3 kg for every 1-unit increment in log TSH concentrations in women ($P < .001$) and by 1.1 kg in men ($P = .002$); the sex interaction term was of borderline significance ($P = .049$). When the analyses were limited to those with a TSH concentration of less than 5 mIU/L at follow-up, resulting in the exclusion of an additional 15 men and 16 women, results were minimally attenuated: weight increased by 1.9 kg per every 1-unit increase in log TSH concentration ($P < .001$) in women and by 1.0 kg per every 1-unit increase in log TSH concentration ($P = .007$) in men. In the quartile-based analysis, there was a strong, graded positive relation between ΔTSH concentration and weight change (**Figure 3**); the mean weight change increased from 0.5 kg in the lowest quartile of ΔTSH concentration to 2.3 kg in the highest quartile ($P < .001$ for trend) in women, and from 0.4 kg in the lowest quartile of ΔTSH concentration to 1.3 kg in the highest quartile ($P = .007$ for trend) in men. When serum TSH concentration at follow-up was limited to a serum TSH concentration of less than 5.0 mIU/L, the results were not materially changed (data not shown).

Table. Characteristics of Study Sample at Baseline and Follow-up[a]

Characteristic	Women (n = 1117)	Men (n = 1290)
Baseline		
Age, y	48 (10)	48 (10)
Weight, kg	66.6 (13.6)	84.1 (12.5)
BMI	25.3 (5.1)	27.2 (3.6)
Serum TSH concentration, mIU/L[b]	1.91 (0.88)	1.70 (0.79)
Current smoking, %	30	28
Postmenopausal, %	50	
Follow-up		
Weight, kg	68.0 (14.5)	85.1 (13.0)
Weight change, kg	1.5 (5.6)	1.0 (5.0)
BMI	26.0 (5.4)	27.7 (3.8)
Serum TSH concentration, mIU/L	1.96 (1.04)	1.83 (0.96)
Serum TSH change, mIU/L[c]	0.06 (0.86)	0.13 (0.81)
Current smoking, %	25	24
Postmenopausal, %	59	

Abbreviations: BMI, body mass index (calculated as weight in kilograms divided by height in meters squared); TSH, thyrotropin.
[a] Unless otherwise indicated, data are expressed as mean (SD).
[b] The 25th, 50th, and 75th percentiles of serum TSH concentrations among women were 1.21, 1.76, and 2.44 mIU/L, respectively; among men, 1.13, 1.53, and 2.09 mIU/L, respectively.
[c] The 25th, 50th, and 75th percentiles of change in serum TSH concentrations among women were –0.34, 0.02, and 0.42 mIU/L, respectively; among men, –0.29, 0.09, and 0.46 mIU/L, respectively.

UPPER DECILE OF WEIGHT CHANGE OVER TIME AND SERUM TSH CONCENTRATION AT FOLLOW-UP

In a secondary analysis of participants at follow-up, including those with abnormal serum TSH values who were not taking thyroid hormone therapy (n = 2632), we compared the participants in the upper decile of weight change (from baseline to follow-up) with the rest of the group. In women, those in the upper decile of weight change (n = 130) gained a mean of 12.0 (range, 7.7-48.2) kg; the mean serum TSH concentration in this group at follow-up was 2.86 mIU/L, compared with 2.41 mIU/L in the rest of the sample. After adjustment for age, the dif-

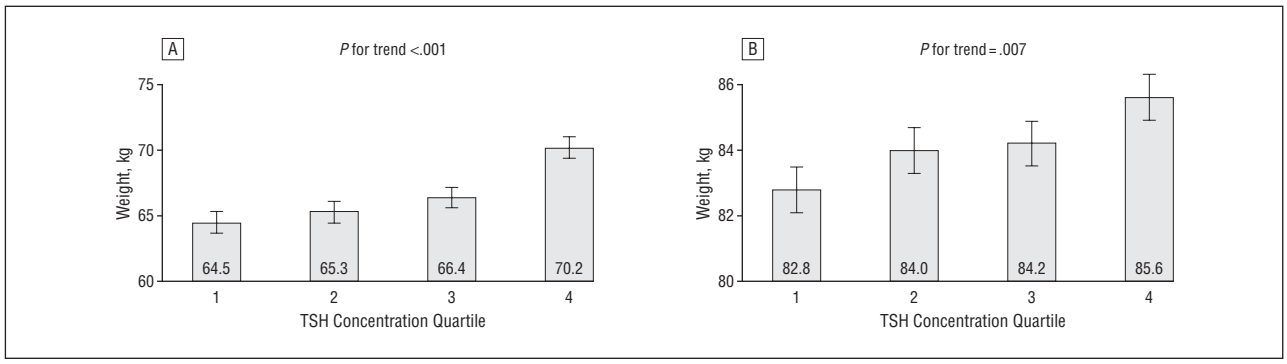


Figure 1. Mean (SE) multivariable-adjusted body weight according to the quartile of serum thyrotropin (TSH) concentration at baseline among women (A) and men (B), adjusted for age, smoking status, and menopausal status (in women). Error bars represent standard error. The quartiles of TSH concentrations for men and women are given in the footnote to the Table.

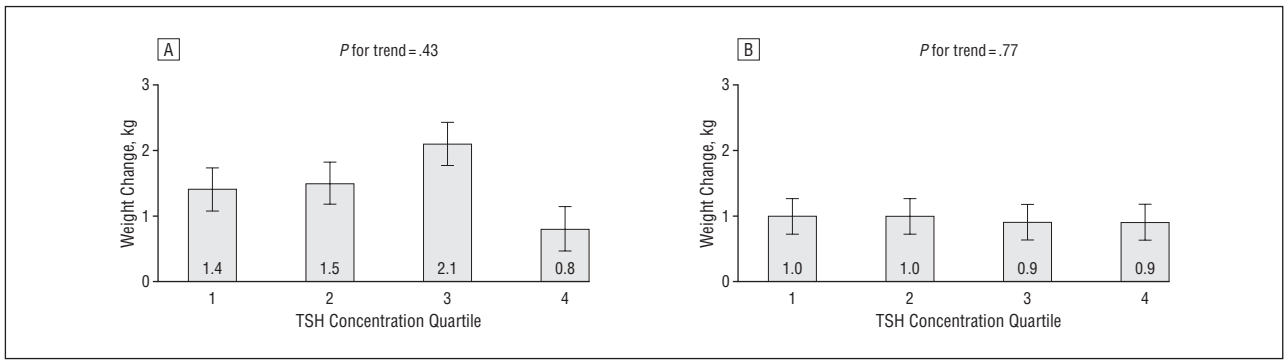


Figure 2. Mean (SE) multivariable-adjusted mean change in body weight according to the quartile of baseline serum thyrotropin (TSH) concentration among women (A) and men (B), adjusted for age, smoking status, menopausal status (in women), and baseline weight. Error bars represent standard error. The quartiles of TSH concentrations for men and women are given in the footnote to the Table.

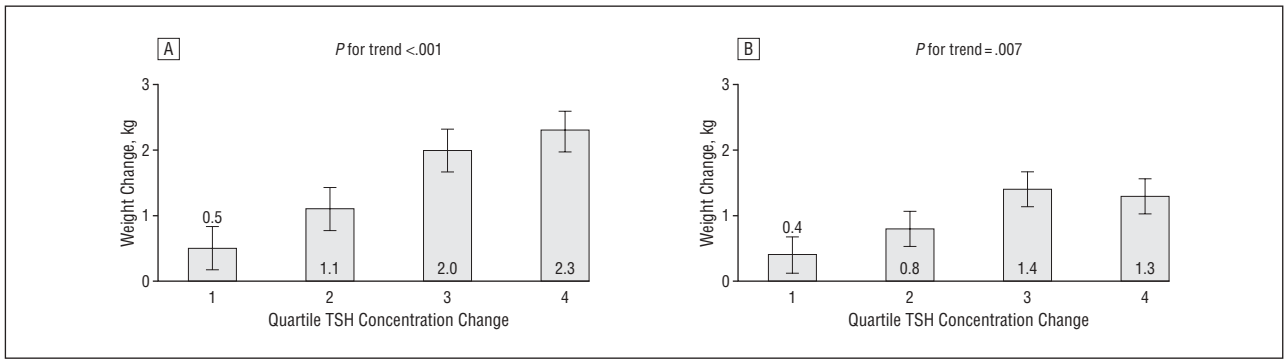


Figure 3. Mean (SE) multivariable-adjusted mean change in body weight according to quartile of change in serum thyrotropin (TSH) concentration among women (A) and men (B), adjusted for age, smoking status, menopausal status (in women), and baseline weight. Error bars represent standard error. The quartiles of TSH concentrations for men and women are given in the footnote to the Table.

ference in log serum TSH concentration was higher among women who gained more weight ($P = .02$). More women in the highest decile of weight change had elevated TSH concentrations at follow-up (serum TSH concentration, >5.0 mIU/L) compared with the rest of the sample (9.2% vs 5.7%), although this difference was of borderline statistical significance ($P = .06$). In men, the mean serum TSH concentration among those in the highest decile of weight change was 1.87 mIU/L (mean weight change, 9.6 [range, 6.8-21.8] kg) compared with 2.04 mIU/L in the rest of the sample ($P = .64$). In men, 2.2% of those in the highest decile of weight change had a serum TSH concentration higher than 5.0 mIU/L at follow-up compared with 2.7% of the remainder of the sample ($P = .80$).

COMMENT

PRINCIPAL FINDINGS

In our community-based sample of participants with thyroid function within the reference range, we observed that the baseline serum TSH concentration was strongly and linearly associated with cross-sectional weight in women and men. During 3.5 years of follow-up, baseline serum

TSH concentrations were not associated with change in weight. However, change in serum TSH concentrations over time (within the reference range) was strongly and linearly associated with weight gain. Models analyzing TSH concentration as a continuous variable yielded results consistent with quartile-based models.

COMPARISON WITH PREVIOUS STUDIES

Data on the association of thyroid function within the range of normal variation and body weight are limited. Patients with overt hypothyroidism lose weight when treated,[18] and patients with hyperthyroidism gain weight when treated.[17,31] Among patients with subclinical hypothyroidism, thyroid hormone therapy has not resulted in significant weight loss.[18,19] However, these studies have been limited by small sample sizes and short follow-up.

There have been few studies of thyroid function and body weight among individuals with normal thyroid function. A Danish study[27] found the following positive cross-sectional association between body mass index and serum TSH concentration: between the lowest and highest quintiles of TSH values, body weight differed by 5.5 kg. A direct comparison of these findings and our results is not possible because these investigators performed sex-pooled analyses and included serum TSH values above and below the accepted reference range. Nonetheless, the results are consistent, lending support to the notion that thyroid function may be an important determinant of body weight. Among 6164 participants from the Tromso Study, women but not men were found to have an association between change in TSH values and body mass index.[24] During 7 years of follow-up, changes in TSH concentration were associated with changes in body weight in sex-pooled analyses; sex-specific analyses were not significant. However, baseline weight was not accounted for in these analyses, a notable limitation given the strong cross-sectional association between weight and TSH concentration. Our data add the important findings that change in serum TSH concentrations over time may be associated with change in body weight, and we extend these findings to women and men after adjustment for several potential important confounders.

POTENTIAL MECHANISMS

There are multiple potential explanations for our findings. There is a well-known association between energy expenditure, thermogenesis, and thyroid function.[32] Lower serum triiodothyronine concentrations (associated with higher serum TSH concentrations) were associated with decreased resting metabolic rate in postobese individuals.[33] In a study of hypothyroid patients receiving long-term treatment with thyroid hormone in whom the dose was varied so as to result in serum TSH concentrations of 0.1 to 10.0 mIU/L, there was a strong inverse association between increases in serum TSH concentration and decreases in resting energy expenditure (by as much as 15%).[34] The authors estimated that this difference could result in an expenditure of approximately 75 to 150 kcal/d, which could result in significant weight gain over time; unfortunately concomitant changes in weight were not reported. Given that low resting energy expenditure is associated with subsequent weight gain,[16] this observation provides a plausible mechanism for the association of change in serum TSH concentration and weight gain over time.

It is intriguing that baseline serum TSH concentrations were strongly associated with baseline body weight, but not with change in body weight over time. Only when the serum TSH concentration increased over time did we find an association with increasing body weight. This finding suggests the possibility of an intrinsic body weight set point that may maintain individuals at a given body weight. Only when there is a slight perturbation in metabolism, as detected by an increase in serum TSH concentration, is an association observed with weight gain.

Our results were stronger in women, although we observed relations similar in directionality in men but with smaller effect sizes. We believe that these sex-related differences in the associations between serum TSH concentration and body weight are probably not related to statistical power; we estimated that we had 92% power to detect in men an effect size similar to that observed in women. Other explanations for our sex-specific findings may be related to the differential effects of thyroid function on body weight and metabolism in women vs men. Fat oxidation is higher in men compared with women[35]; whether this is mediated in part through thyroid function is unknown. The preponderance of thyroid disease in women[29] suggests a differential effect of thyroid pathophysiologic mechanisms in men compared with women.

STRENGTHS AND LIMITATIONS

The strengths of our study include a well-characterized, large community-based sample. We excluded participants with serum TSH values outside the range considered physiologic. We had well-characterized covariate data, enabling us to perform multivariable-adjusted analyses. Study limitations include the observational design, so we cannot infer causality. Indeed, one might question whether weight gain may cause increases in serum TSH concentrations. However, in a study of subjects admitted to a metabolic ward, those who lost 10% of their body weight had a decline in serum TSH concentrations from 3. 1 to 2.4 mIU/L, whereas those who gained 10% of their body weight had no change in serum TSH concentration.[36] Thus, it is unlikely that the increases in serum TSH concentration in our study are due to weight gain alone. An additional limitation of our study is that we did not measure free thyroxine levels. Nonetheless, serum TSH concentrations are generally considered to be the most sensitive marker of thyroid function. In our community-based sample, it is unlikely that other rare causes of hypothyroidism are driving the associations observed. Our sample was nearly entirely white, and our results may not be generalizable to other ethnic groups. Finally, we were unable to account for other covariates known to be associated with body weight and weight change, including diet and physical activity.

CLINICAL IMPLICATIONS

The identification of change in thyroid function as a risk factor for weight gain might help guide research into the identification, prevention, and treatment of individuals at

risk for the development of excess adiposity.[37] Confirmation of our findings in other samples is warranted, and in particular more longitudinal studies are warranted. We attempted to answer the clinical question of whether derangements in thyroid function are present in community-dwelling individuals who gain excess amounts of body weight. In doing so, we found that mean serum TSH concentrations were higher among women in the upper decile of weight change, and that these women were at nonsignificantly increased odds of having serum TSH concentrations higher than 5.0 mIU/L. Therefore, abnormalities in thyroid function may play a small role in significant weight gain among women in an unselected sample.

In conclusion, thyroid function (as assessed by serum TSH concentrations) within the reference range is associated with body weight in both sexes. Our findings raise the possibility that modest increases in serum TSH concentrations within the reference (physiologic) range may be associated with weight gain.

Accepted for Publication: August 12, 2007.
Correspondence: Caroline S. Fox, MD, MPH, National Heart, Lung, and Blood Institute's Framingham Heart Study, 73 Mt Wayte Ave, Ste 2, Framingham, MA 01702 (foxca@nhlbi.nih.gov).
Author Contributions: Dr Fox 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:* Fox, Pencina, D'Agostino, Murabito, Seely, and Vasan. *Acquisition of data:* D'Agostino. *Analysis and interpretation of data:* Fox, Pencina, D'Agostino, Murabito, Seely, Pearce, and Vasan. *Drafting of the manuscript:* Fox and Vasan. *Critical revision of the manuscript for important intellectual content:* Pencina, D'Agostino, Murabito, Seely, Pearce, and Vasan. *Statistical analysis:* Fox, Pencina, and D'Agostino. *Administrative, technical, and material support:* Vasan.
Financial Disclosure: None reported.
Funding/Support: This study was supported by grant N01-HC-25195 from the National Heart, Lung, and Blood Institute's (NHLBI's) Framingham Heart Study. Dr Vasan was supported by grant 2K24 HL04334 from the NHLBI, and Dr Pearce was supported by grant 5K23 DK64611 from the National Institute of Diabetes and Digestive and Kidney Diseases.
Additional Contributions: Joseph Hollowell, MD, MPH, provided the NHANES III data. Robert D. Utiger, MD, provided a critical review of this manuscript. The late Clark T. Sawin, MD, performed the TSH assays.

REFERENCES

1. Mokdad AH, Ford ES, Bowman BA, et al. Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001. *JAMA*. 2003;289(1):76-79.
2. Must A, Spadano J, Coakley EH, Field AE, Colditz G, Dietz WH. The disease burden associated with overweight and obesity. *JAMA*. 1999;282(16):1523-1529.
3. Colditz GA, Willett WC, Rotnitzky A, Manson JE. Weight gain as a risk factor for clinical diabetes mellitus in women. *Ann Intern Med*. 1995;122(7):481-486.
4. Hubert HB, Feinleib M, McNamara PM, Castelli WP. Obesity as an independent risk factor for cardiovascular disease: a 26-year follow-up of participants in the Framingham Heart Study. *Circulation*. 1983;67(5):968-977.
5. Calle EE, Thun MJ, Petrelli JM, Rodriguez C, Heath CW Jr. Body-mass index and mortality in a prospective cohort of US adults. *N Engl J Med*. 1999;341(15):1097-1105.
6. Calle EE, Rodriguez C, Walker-Thurmond K, Thun MJ. Overweight, obesity, and mortality from cancer in a prospectively studied cohort of US adults. *N Engl J Med*. 2003;348(17):1625-1638.
7. Flegal KM, Carroll MD, Kuczmarski RJ, Johnson CL. Overweight and obesity in the United States: prevalence and trends, 1960-1994. *Int J Obes Relat Metab Disord*. 1998;22(1):39-47.
8. Kuczmarski RJ, Flegal KM, Campbell SM, Johnson CL. Increasing prevalence of overweight among US adults: the National Health and Nutrition Examination Surveys, 1960 to 1991. *JAMA*. 1994;272(3):205-211.
9. Zurlo F, Ferraro RT, Fontvielle AM, Rising R, Bogardus C, Ravussin E. Spontaneous physical activity and obesity: cross-sectional and longitudinal studies in Pima Indians. *Am J Physiol*. 1992;263(2, pt 1):E296-E300.
10. Tataranni PA, Harper IT, Snitker S, et al. Body weight gain in free-living Pima Indians: effect of energy intake vs expenditure. *Int J Obes Relat Metab Disord*. 2003;27(12):1578-1583.
11. Rissanen AM, Heliovaara M, Knekt P, Reunanen A, Aromaa A. Determinants of weight gain and overweight in adult Finns. *Eur J Clin Nutr*. 1991;45(9):419-430.
12. Engström G, Hedblad B, Stavenow L, Lind P, Janzon L, Lindgärde F. Inflammation-sensitive plasma proteins are associated with future weight gain. *Diabetes*. 2003;52(8):2097-2101.
13. Sammel MD, Grisso JA, Freeman EW, et al. Weight gain among women in the late reproductive years. *Fam Pract*. 2003;20(4):401-409.
14. Snyder EE, Walts B, Perusse L, et al. The human obesity gene map: the 2003 update. *Obes Res*. 2004;12(3):369-439.
15. Kershaw EE, Flier JS. Adipose tissue as an endocrine organ. *J Clin Endocrinol Metab*. 2004;89(6):2548-2556.
16. Ravussin E, Lillioja S, Knowler WC, et al. Reduced rate of energy expenditure as a risk factor for body-weight gain. *N Engl J Med*. 1988;318(8):467-472.
17. Dale J, Daykin J, Holder R, Sheppard MC, Franklyn JA. Weight gain following treatment of hyperthyroidism. *Clin Endocrinol (Oxf)*. 2001;55(2):233-239.
18. Tzotzas T, Krassas GE, Konstantinidis T, Bougoulia M. Changes in lipoprotein(a) levels in overt and subclinical hypothyroidism before and during treatment. *Thyroid*. 2000;10(9):803-808.
19. Kong WM, Sheikh MH, Lumb PJ, et al. A 6-month randomized trial of thyroxine treatment in women with mild subclinical hypothyroidism [published corrections appear in *Am J Med*. 2002;113(3):264 and *Am J Med*. 2002;113(5):442]. *Am J Med*. 2002;112(5):348-354.
20. Michalaki MA, Vagenakis AG, Leonardou AS, et al. Thyroid function in humans with morbid obesity. *Thyroid*. 2006;16(1):73-78.
21. Iacobellis G, Ribaudo MC, Zappaterreno A, Iannucci CV, Leonetti F. Relationship of thyroid function with body mass index, leptin, insulin sensitivity and adiponectin in euthyroid obese women. *Clin Endocrinol (Oxf)*. 2005;62(4):487-491.
22. Manji N, Boelaert K, Sheppard MC, Holder RL, Gough SC, Franklyn JA. Lack of association between serum TSH or free T4 and body mass index in euthyroid subjects. *Clin Endocrinol (Oxf)*. 2006;64(2):125-128.
23. Knudsen N, Laurberg P, Rasmussen LB, et al. Small differences in thyroid function may be important for body mass index and the occurrence of obesity in the population. *J Clin Endocrinol Metab*. 2005;90(7):4019-4024.
24. Nyrnes A, Jorde R, Sundsfjord J. Serum TSH is positively associated with BMI. *Int J Obes (Lond)*. 2006;30(1):100-105.
25. Vasan RS, Pencina MJ, Cobain M, Freiberg MS, D'Agostino RB. Estimated risks for developing obesity in the Framingham Heart Study. *Ann Intern Med*. 2005;143(7):473-480.
26. Dawber TR, Kannel WB, Lyell LP. An approach to longitudinal studies in a community: the Framingham Heart Study. *Ann N Y Acad Sci*. 1963;107:539-556.
27. Shurtleff D. Some characteristics related to the incidence of cardiovascular disease and death: Framingham Study, 18-year follow-up, section 30. In: Kannel WB, Fordon T, eds. *The Framingham Study: An Epidemiological Investigation of Cardiovascular Disease*. Washington, DC: Dept of Health, Education, and Welfare; 1973. DHEW publication (NIH) 74-599.
28. Sawin CT, Geller A, Wolf PA, et al. Low serum thyrotropin concentrations as a risk factor for atrial fibrillation in older persons. *N Engl J Med*. 1994;331(19):1249-1252.
29. Hollowell JG, Staehling NW, Flanders WD, et al. Serum TSH, T_4, and thyroid antibodies in the United States population (1988 to 1994): National Health and Nutrition Examination Survey (NHANES III). *J Clin Endocrinol Metab*. 2002;87(2):489-499.
30. SAS Institute Inc. *SAS/STAT User's Guide, Version 9.1*. Cary, NC: SAS Institute Inc; 2004.
31. Brunova J, Bruna J, Joubert G, Koning M. Weight gain in patients after therapy for hyperthyroidism. *S Afr Med J*. 2003;93(7):529-531.
32. Silva JE. The thermogenic effect of thyroid hormone and its clinical implications. *Ann Intern Med*. 2003;139(3):205-213.
33. Astrup A, Buemann B, Toubro S, Ranneries C, Raben A. Low resting metabolic rate in subjects predisposed to obesity: a role for thyroid status. *Am J Clin Nutr*. 1996;63(6):879-883.
34. al-Adsani H, Hoffer LJ, Silva JE. Resting energy expenditure is sensitive to small dose changes in patients on chronic thyroid hormone replacement. *J Clin Endocrinol Metab*. 1997;82(4):1118-1125.
35. Toth MJ, Gardner AW, Arciero PJ, Calles-Escandon J, Poehlman ET. Gender differences in fat oxidation and sympathetic nervous system activity at rest and during submaximal exercise in older individuals. *Clin Sci (Lond)*. 1998;95(1):59-66.
36. Rosenbaum M, Hirsch J, Murphy E, Leibel RL. Effects of changes in body weight on carbohydrate metabolism, catecholamine excretion, and thyroid function. *Am J Clin Nutr*. 2000;71(6):1421-1432.
37. Krotkiewski M. Thyroid hormones in the pathogenesis and treatment of obesity. *Eur J Pharmacol*. 2002;440(2-3):85-98.

National Cancer Institute Patient Navigation Research Program

Methods, Protocol, and Measures

Karen M. Freund, MD, MPH[1]
Tracy A. Battaglia, MD, MPH[1]
Elizabeth Calhoun, PhD[2]
Donald J. Dudley, MD[3]
Kevin Fiscella, MD MPH[4]
Electra Paskett, PhD[5]
Peter C. Raich, MD[6]
Richard G. Roetzheim, MD[7]
The Patient Navigation Research Program Group

[1] Women's Health Unit, Department of Medicine, and Women's Health Interdisciplinary Research Center, Boston University School of Medicine, Boston, Massachusetts.

[2] Division of Health Policy Administration, School of Public Health, University of Illinois at Chicago, Chicago, Illinois.

[3] Department of Obstetrics and Gynecology, University of Texas Health Science Center, San Antonio, Texas.

[4] Department of Family Medicine and Community and Preventive Medicine and Division of Oncology, James P. Wilmont Cancer Center, University of Rochester, School of Medicine, Rochester, New York.

[5] Division of Epidemiology, Comprehensive Cancer Center, Ohio State University, Columbus, Ohio.

[6] Division of Hematology/Oncology, Department of Medicine, Denver Health and Hospital Authority, Denver, Colorado.

[7] Department of Family Medicine, University of South Florida, Division of Health Outcomes and Behaviors, H. Lee Moffitt Cancer Center and Research Institute, Tampa, Florida.

Supported by National Institutes of Health (NIH) Grants U01 CA116892, U01 CA 117281, U01CA116903, U01CA116937, U01CA116924, U01CA116885, U01CA116875, U01CA116925, and American Cancer Society grant #SIRSG-05-253-01.

BACKGROUND. Patient, provider, and systems barriers contribute to delays in cancer care, a lower quality of care, and poorer outcomes in vulnerable populations, including low-income, underinsured, and racial/ethnic minority populations. Patient navigation is emerging as an intervention to address this problem, but navigation requires a clear definition and a rigorous testing of its effectiveness. Pilot programs have provided some evidence of benefit, but have been limited by evaluation of single-site interventions and varying definitions of navigation. To overcome these limitations, a 9-site National Cancer Institute Patient Navigation Research Program (PNRP) was initiated.

METHODS. The PNRP is charged with designing, implementing, and evaluating a generalizable patient navigation program targeting vulnerable populations. Through a formal committee structure, the PNRP has developed a definition of patient navigation and metrics to assess the process and outcomes of patient navigation in diverse settings, compared with concurrent continuous control groups.

RESULTS. The PNRP defines patient navigation as support and guidance offered to vulnerable persons with abnormal cancer screening or a cancer diagnosis, with the goal of overcoming barriers to timely, quality care. Primary outcomes of the PNRP are 1) time to diagnostic resolution; 2) time to initiation of cancer treatment; 3) patient satisfaction with care; and 4) cost effectiveness, for breast, cervical, colon/rectum, and/or prostate cancer.

CONCLUSIONS. The metrics to assess the processes and outcomes of patient navigation have been developed for the NCI-sponsored PNRP. If the metrics are found to be valid and reliable, they may prove useful to other investigators. ***Cancer* 2008;113:3391–9.** © *2008 American Cancer Society.*

KEYWORDS: breast cancer, cervical cancer, colorectal cancer, prostate cancer, navigation, case management, minority groups, medically underserved areas, vulnerable populations.

The following are members of the Patient Navigation Research Group: Charles L. Bennett, MD (Northwestern University Feinberg School of Medicine; Jack A. Clark, PhD (Boston University School of Public Health); Roland Garcia, PhD (Center to Reduce Cancer Health Disparities, National Cancer Institute) Feinberg School of Meicine; Amanda Greene, PhD (NOVA Research); Steven R. Patierno, PhD (George Washington University Cancer Institute); and Victoria Warren-Mears, PhD, RD, LD (Northwest Tribal Epidemiology Center).

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, National Cancer Institute.

Address for reprints: Karen M. Freund, MD, MPH, Boston University School of Medicine, 801 Massachusetts Avenue, Crosstown 470, Boston, MA 02118; Fax: (617) 638-8096; E-mail: Karen.freund@bmc.org

Received April 22, 2008; revision received July 14, 2008; accepted July 21, 2008.

DOI 10.1002/cncr.23960
Published online 24 October 2008 in Wiley InterScience (www.interscience.wiley.com).

In recent decades, advances in screening and treatment have resulted in improved cancer outcomes.[1] However, disparities in cancer outcomes according to race and income continue.[2-4] A recent focus in cancer research has been to understand the social, economic, cultural, behavioral, and systems barriers to receiving comprehensive cancer care in a timely fashion and to eliminate these persistent disparities.[5-7] Patient navigation represents 1 proposed remedy for disparities in cancer outcomes by intervening to address these barriers to care.[8] Several uncontrolled studies and small single-site trials have suggested that patient navigation may improve cancer outcomes.[9-11]

Patient navigation has been defined as the logistic and emotional support needed to achieve the completion of diagnostic and treatment care. Individuals previously identified as case managers, patient advocates, community health workers, and schedule coordinators are now being placed under the umbrella of 'patient navigation.' Although the concepts of patient navigation can be used for multiple chronic and acute diseases,[12] the lack of a common nomenclature with clearly defined job responsibilities makes comparison of different navigator models difficult.

Before patient navigation can be extended as a standard of cancer care, empiric evidence of its benefit and cost-effectiveness must be demonstrated. To our knowledge to date, there are no accepted measures of either the navigation process or its clinical and economic outcomes. The development and dissemination of process and outcome measures will allow communities and researchers to evaluate the results of these programs.

To our knowledge the Patient Navigation Research Program (PNRP) sponsored by the National Cancer Institute's (NCI) Center to Reduce Cancer Health Disparities (CRHCD) is the first multicenter program to examine the role and benefits of patient navigation. To achieve this aim, the Steering Committee of the PNRP developed a definition of patient navigation and a series of common measures to assess outcomes of care with patient navigation. We present herein the definitions and measures developed to assess the benefits of patient navigation.

MATERIALS AND METHODS

Overview of Program

Funded through the NCI with additional support from the American Cancer Society and the Avon Foundation, the PNRP is a cooperative effort of 9 sites across the US. Target communities include racial and ethnic minorities and those of low socioeconomic status who have either abnormal cancer screening or an incident diagnosis of breast, cervical, colorectal, or prostate cancer. Investigators in each site will assess the outcomes in a group of patients receiving patient navigation, compared with a concurrent control group without navigation.

Definition of Navigation and Role of Navigators

The working definition of patient navigation was provided by the NCI's CRHCD in their request for applications.[13] In this definition, patient navigation refers to the support and guidance offered to persons with abnormal cancer screening or a new cancer diagnosis in accessing the cancer care system; overcoming barriers; and facilitating 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 racial and ethnic minorities and those from low-income populations. Furthermore, patient navigation targets specific timepoints in the cancer care continuum; we operationally define patient navigation as starting at the time of an abnormal screening result and ending at the determination that the screening test was a false-positive or, for those individuals with a new cancer diagnosis, continuing through the completion of cancer treatment. The goal of patient navigation is to facilitate timely access to quality cancer care that meets cultural needs and standards of care for all patients.

Examples of navigation services include: arranging various forms of financial support, arranging for transportation to and childcare during scheduled appointments, identifying and scheduling appointments with culturally sensitive caregivers, coordinating care among providers, arranging for interpreter services, ensuring coordination of services among medical personnel, ensuring that medical records are available at each scheduled appointment, and providing other services to overcome access barriers encountered during the cancer care process including linkage to community resources. Navigators work to address health literacy and to train patients to advocate for themselves in the healthcare system. They are also trained to provide emotional support to patients during this stressful period. Navigators may also identify systems issues that serve as barriers to many patients, and work toward reducing the complexity to the patient of the multidisciplinary approach to care.

The concept of patient navigation is based on the care management or case management model, which has 4 components.[14] The first is case identification, which is a systematic approach to the identification of those individuals with abnormal cancer screening in need of follow-up care or incident cancers. The second is identifying individual barriers to

TABLE 1
Cancer Type and Populations Studied by National Site–NCI Patient Navigation Program

	Cancer Type				Populations Addressed			
National Site	Breast	Cervical	Colorectal	Prostate	African American	Hispanic	Asian	American Indian/ Native Alaskan
Boston University Medical Center	X	X			X	X	X	
Denver Health and Hospital Authority	X		X	X	X	X	X	X
H. Lee Moffitt Cancer Center	X		X		X	X		
Northwest Portland Area Indian Health Board	X	X	X	X				X
Northwestern University			X	X	X	X		
University of Illinois/Access Community Health Network	X	X			X	X		
George Washington University	X				X	X		
Ohio State University	X	X	X		X	X		
University of Rochester School of Medicine and Dentistry	X		X		X			
University of Texas Health Science Center	X	X			X	X		

NCI indicates National Cancer Institute.

receiving care. Navigators contact patients and elicit information regarding the barriers to completion of recommended care. The third is developing an individualized plan to address the barriers that are identified. The fourth is tracking, which is a systematic method of following each case through resolution of the problem. In the case of cancer navigation, this is to resolution of a diagnostic evaluation when a benign condition is diagnosed or follow-up to the completion of primary therapy when a cancer or premalignant condition is diagnosed.

The navigator will focus on assisting patients and coordinating care of the patients among providers, community, and the patients and their families. Given that patient navigators are working primarily with racial/ethnic minority and low-income patients, cultural competence is a key feature. Cultural and linguistic competence is a set of congruent behaviors, attitudes, and policies that enable effective work in cross-cultural situations.[15]

PNRP sites vary in the prior training, skill sets, and educational background of navigators and include lay community peers, health educators and advocates, medical assistants, social workers, and nurses. The study has set a minimum requirement of a high school diploma or General Education Diploma. In an effort to achieve a core set of knowledge, skills, and competencies across navigators, a standardized training has been developed. The curriculum focuses on basic information regarding cancer and its diagnosis and treatment, professionalism, understanding barriers to care, communication skills, cultural competency, ethical conduct of human subjects research, and developing a local network of resources to support patients.[16]

Cancers of Interest

The PNRP program chose breast, cervical, colorectal, and prostate cancer for several reasons. Each is prevalent, particularly in low-income populations, accounts for significant morbidity and mortality, and there exists evidence of racial and ethnic outcome disparities.[2] For these cancers, there is a commonly used screening test and evidence of better outcomes with earlier-stage disease for cervical, breast, and colorectal cancer,[17,18] with trials under way to assess the benefits of prostate cancer screening.[19] Each of the 9 sites is addressing 1 or more of the cancers in specific underserved populations (Table 1).

Definition of Metrics and Methods

Key variables necessary to answer the primary outcome questions were required of all sites and form the minimal or 'common' dataset for all sites to collect. Multiple secondary analyses and subquestions have emerged. As each of these questions arose, common metrics were chosen for these additional 'optimal' elements, so that the sites collecting this additional information could pool their data for analyses of secondary outcomes and research questions. Metrics were developed by use of those guidelines that exist within the medical literature or by consensus of the steering committee. The National Comprehensive Cancer Network (NCCN) guidelines formed the major focus of the clinical guidelines on management for both screening abnormalities and diagnostic management decisions.[20-23] The steering committee also reviews relevant changes in guidelines during the course of the study.[24]

TABLE 2
Patient Navigation Research Program Eligibility Criteria

Cancer Site	Test	Abnormality
Breast screening abnormality	Clinical breast examination	• Breast mass • Clinical finding suspicious for cancer
	Screening mammogram	• BIRADS 0, 3, 4, 5
	Screening ultrasound	• BIRADS 0, 3, 4, 5
	Screening MRI	• BIRADS 0, 3, 4, 5
Breast cancer	Pathology	• DCIS • Invasive cancer
Cervical screening abnormality	Cytology	• LGSIL* • HGSIL • ASCUS, HPV positive* • ASCUS, no HPV testing* • AGUS
	Clinical examination	• Suspicious abnormality
Cervical cancer and precancerous lesions	Pathology	• CIN2 • CIN3 • CIS • Invasive cervical cancer
Colorectal screening abnormality	Clinical history and examination	• Blood in stool or rectal bleeding in patient aged ≥30 y • Rectal mass
	FOBT	• Positive
	Sigmoidoscopy or colonoscopy	• Polyp • Space–occupying lesion
	Double-cntrast barium enema	• Space–occupying lesion
	Virtual colonoscopy	• Space–occupying lesion
Colorectal cancer	Pathology	• CIS • Invasive cancer
Prostate screening abnormality	Clinical examination	• Prostate induration • Prostate nodule • Prostate asymmetry
	PSA	• Abnormal PSA • Abnormal PSA velocity
Prostate cancer	Pathology	• PIN or CIS • Invasive cancer

BIRADS indicates Breast Imaging Reporting and Data System; MRI, magnetic resonance imaging; DCIS, ductal carcinoma in situ; LGSIL, low-grade squamous intraepithelial lesion; HGSIL, high-grade squamous intrapeithelial lesion; ASCUS, atypical cells of undetermined significance; HPV, human papillomavirus; AGUS, atypical glandular cells of undetermined significance; CIN2, cervical intraepithelial lesion of type 2; CIN3, CIN of type 3; CIS, carcinoma is situ; FOBT, fecal occult blood test; PSA, prostate-specific antigen; PIN, prostatic intraepithelial neoplasia.
*For women aged ≥21 years.

Eligibility and Exclusion Criteria

Table 2 lists the screening abnormalities and diagnostic categories eligible for inclusion into the study. For each cancer, abnormal findings on screening studies that require additional testing are included. For each disease, clinical findings suspicious for cancer, for example, a breast mass, rectal bleeding, or suspicious cervical lesions that result in referral to a disease specialist, will also serve as entry criteria. Lastly, a patient can enroll if presenting with a cancer diagnosis without prior treatment.

Exclusion criteria included a prior history of cancer other than nonmelanoma skin cancer, because patients who have already experienced the multidisciplinary complexity of cancer treatment may be more likely to be able to navigate the system. Patients with prior abnormal cancer screenings but without a cancer diagnosis are eligible. Patients who have received patient navigation for a cancer screening abnormality are excluded because the benefits of their prior navigation may confound the current intervention. Patients who have experience with case coordination for another disease process, such as mental health or diabetes care management, are not excluded; however, information regarding their prior care coordination will be collected. Women who are pregnant at study entry are excluded because delays in care influenced by pregnancy status, such as postponing cervical biopsy for cervical lesions until after delivery, will confound comparisons of the course of follow-up care.

Methods of Case and Control Allocation

The issue of allocating subjects to intervention versus control arms for an intervention that appears intuitively beneficial has ethical considerations. However, none of the sites in the program had pre-existing navigation services, nor were there other resources available for navigation; therefore, the control groups were not denied a service that would otherwise be available to them. Each site has developed a method of allocation of cases and controls to address scientific rigor and logistic needs of working with community partners, within the context of these ethical concerns, and all were approved by their Institutional Review Board (IRB). Several sites will conduct randomized clinical trials, with randomization at the individual level at each site. Some sites have reported community concerns about not providing all eligible patients the opportunity for entry into the navigation arm, and several sites have expressed concern about contamination when attempting individual randomization. Several sites that are recruiting from multiple community healthcare centers have opted to randomize each clinical site to either case or control status as a way to address the above concerns. One site, in response to community concerns and buy-in for the project, has allocated each site as a navigation site for 1 type of cancer and control status for another cancer. Some sites have provided a minimal education intervention to the control arm, to facilitate buy-in from providers and subjects.

Primary Outcomes

Timeliness of diagnosis

Four primary outcomes were selected that are clinically relevant and for which disparities in care among racial and ethnic minorities and/or low-income individuals have been documented (Table 3). The ultimate outcome of an effective cancer intervention is reduction in morbidity and mortality. Delays in follow-up of abnormal cancer screening can often result in increased patient morbidity and mortality.[25,26] Our first outcome measures will be the intermediate outcome of time to completion of diagnostic evaluation, because we do not have the power or the longitudinal design in this study to measure cancer-specific mortality. Screening abnormality is defined in our study as the date that the screening test was conducted (eg, the date of an abnormal clinical breast examination or the date that a prostate-specific antigen or cervical cytology specimen was collected). We chose this definition because the date of report of the abnormal result, date of physician notification, or date of patient notification can reflect delays. Diagnostic resolution is defined as completion of the diagnostic test that results in a diagnosis or clinical evaluation that determines that no further evaluation is indicated. For example, a colonoscopy with biopsy confirming a malignant polyp or a colonoscopy in which no malignant lesion is identified would both serve as a diagnostic resolution.

TABLE 3
Outcomes to Evaluate Patient Navigation

Primary Outcomes	Secondary Questions
Time to completion of diagnosis	Time to completion of therapy
Time to initiation of primary therapy	Quality of care
Patient satisfaction and quality of life	Navigator characteristics
Cost-effectiveness	Task and social network analysis

Timeliness of cancer treatment

Subjects can enter the cancer treatment phase of the study either from the diagnostic phase, when a cancer diagnosis is established, or as a new subject with an untreated cancer. We will record the date at which the diagnostic test was performed that established the cancer diagnosis and the date at which cancer treatment was initiated (eg, the date of biopsy of a polyp and the date on which a partial colectomy is performed).

Quality of life and patient satisfaction

The news of a positive screening test for cancer is likely to cause immediate quality of life changes, including emotional distress. We selected the Impact of Events Scale (IES) as a common validated measure of health-related quality of life,[27] which is widely used in cancer studies. The IES addresses the distress, intrusive thoughts, and misgivings precipitated by the troublesome event of an abnormal screening result or a diagnosis of cancer and can also be adapted to refer to a specific screening test result or diagnosis of cancer without altering its meaning or measurement properties. The IES will be used to collect data at both the postscreening follow-up and postdiagnosis treatment phases of a patient's experiences.

To measure patient self-efficacy in dealing with cancer and related health services, we chose the newly developed Communication and Attitudinal Self-Efficacy Scale (CASE) measure.[28] CASE was validated in a diverse population of general oncology patients. The CASE has 2 forms: generic and cancer-focused. The former assesses self-efficacy in dealing with healthcare in ways that are relevant to follow-up after a positive screening. The latter form specifically assesses self-efficacy in dealing with the healthcare challenges after a cancer diagnosis.

Although there are many measures of patient satisfaction, we found none that was specifically relevant to the expected experiences of care and the perceptions of the study participants experiencing navigation. We are developing a navigation-focused measure that will assess satisfaction with aspects of care in which navigation may be expected to have an impact. To do this, we have adapted domains and items from existing measures and developed new items, based on the combined expertise of the 9 research teams. The resulting 29-item instrument addresses 3 major domains of patient satisfaction with clinical encounters: interpersonal process, outcomes, and structural/access issues. We are currently conducting psychometric evaluation to validate this new instrument, with data from the first 500 subjects surveyed by the 9 sites.

In addition to satisfaction with care, we are developing a measure of satisfaction with the navigator, consisting of 2 scales. One scale includes 26 items that assess a subject's perceptions of the effectiveness of his or her navigator's efforts in overcoming specific barriers to care after a positive screening test or diagnosis, such as scheduling appointments, completing forms, and dealing with child care issues. These items correspond to the content of training that navigators receive and the specific list of barriers to care they are trained to investigate and address. The second scale consists of 9 items that assess the patient's subjective satisfaction with the interpersonal relationship with the navigator. This new scale will also be subjected to validation with data from initial subjects.

We determined that the time frame for measuring patient satisfaction and quality of life would be within 3 months of the completion of a diagnostic evaluation for patients with abnormal screening and within 3 months of the initiation of cancer treatment for patients with cancer. This time frame was chosen to reflect the logistical issues of reaching patients to complete the inventory of items, while remaining within a time period in which the impact of events and their satisfaction with the care they received would remain current issues.

Cost-effectiveness analysis

A cost-effectiveness analysis will compare the costs of care using a patient navigation model with usual care against estimates of quality-of-life years. A societal view of costs will be used to include estimates of start-up costs for navigation programs, training costs, and fixed and variables cost of the program, as well as healthcare expenditures and patient out-of-pocket costs associated with their medical care. Patient utilities will be derived from a subset of patients in the navigated and control arm using generic multiattribute utility instruments. The timeframe of the study is too short to directly measure survival time; this will be inferred from the distribution of stage at diagnosis in each group.

Secondary Outcomes

Completion of therapy

Many sites will collect data regarding therapy completion and allow us to examine whether navigation improves rates of completion of radiotherapy and chemotherapy. Because current data suggest that the timing of therapies may play a role in effectiveness, and delays or incomplete therapy may impede effectiveness,[29,30] these data will address this potential benefit of navigation.

Quality of care

By collecting details regarding the staging of each cancer diagnosis and the therapies completed by patients, and using evidence-based guidelines on therapeutic choices,[20-23] we can make some assessments of quality of care. For example, we can examine proportions of eligible women with estrogen receptor-positive breast tumors in the navigated and control arms who are offered and who receive hormonal therapy.

Process of patient navigation

Understanding the content of the work of navigation is critical to document the exact nature of the intervention (ie, the work activities of the navigators). Also, in order for other studies to compare their findings with the PNRP, there is a need for common metrics to measure navigation. Currently, no such metrics exist to our knowledge. We propose the following methods for other researchers to implement when evaluating patient navigation programs.

We have developed a common patient log for navigators to complete to document their work with patients. The log is based on each direct contact with the patient and the activities performed on behalf of the patient. The nature of each patient contact (eg, by telephone, E-mail, or in person, and at what site) and the duration of the encounter are recorded. Navigators will document barriers to care from a predefined list and actions taken by navigators to address these barriers. Variables also include the navigator estimate of total time to address each case. These variables will allow us to compare the barriers to care across sites, identify barriers that are not overcome, and identify which actions are asso-

TABLE 4
Data Elements in Navigator Tracking Log

Date of encounter
Length of time
Length of direct encounter time with the patient (categorical by 15-min intervals to >90 min)
Total navigation time to complete navigation activities outside the time spent with the patient (categorical by 15-min intervals to >240 min)
Type of encounter
In-person home visit
In-person clinic/hospital visit at site at where navigator is based
In-person at hospital or clinical site other than where navigator is based
In-person at nonclinical site (eg, social service agency, support group)
Phone call with patient
Written message to patient (letter, E-mail)
Communication with family, friends, church
Communication with medical staff
Communication with nonmedical staff
Communication with resource staff
Barriers
Transportation
Housing
Social/practical support
Language/interpreter
Literacy
Childcare issues
Adult care
Location of healthcare facility
Insurance issues
Financial problems
Employment issues
Communication concerns with medical personnel
Fear
Medical and mental health comorbidity
Patient disability
Out of town/country
Perceptions/beliefs about tests/treatment
System problems with scheduling care
Attitudes towards providers
Other barrier name
No barriers identified
Actions
Referrals/direct contact
Accompaniment
Arrangements
Support
Records/record keeping
Education
Scheduling appointments
Directly contacting family
Systems proactive navigation
No actions taken
Other

ciated with improvements in outcomes with navigation (Table 4).

We are collecting information regarding the navigators themselves, including prior training and experience and their personal and family experience with cancer. The demographic data collected about the navigators are analogous to those collected on patients, including race and ethnicity, sex, language, health insurance, housing, and family dependents. These variables will allow us to assess whether specific characteristics of navigators are associated with successful navigation and whether congruence on demographic characteristics between navigators and patients promotes improved care.

We plan direct observation of the activities of navigators to assess the content of their activities. Each navigator will be assessed twice yearly on an 8-point competency checklist to ensure that minimum standards are met across the 9 sites. We hypothesize that the effectiveness of the patient navigator is related to networking of resources available to the navigator to assist in care. This requires that the navigator have access to a network of resources to support the patients' needs. Therefore, we are developing a new structured observation protocol, using concepts from task and social network analysis,[31-33] that will assess through direct observation of the navigators 2 dimensions: the individuals and organizational entities with whom the navigator interacts and the type of task performed out in support of the navigated patient.

Each of the 9 research sites is conducting the intervention in multiple healthcare settings. Information is collected annually concerning each clinical care center: geographic location, annual clinical volume, race and ethnicity of patients seen, and onsite services related to cancer screening and diagnosis. Other optional variables are collected by some of the sites for subset analyses. These include comorbidity using the Charlson comorbidity score[34] and a family history of cancer. Literacy is assessed using the Rapid Estimate of Adult Literacy in Medicine (REALM)[35] or by self-report of problems in reading instructions and health information.[36]

DISCUSSION

Patient navigation was a term first used to describe the case management of patients in need of cancer screening or with cancer screening abnormalities.[8] This term is now being widely used to describe a broad array of roles and functions, from traditional administrative assistant positions, community outreach workers, social workers, nurses, and patient advocates.[9] The diversity of job and role descriptions, coupled with little available data regarding the outcomes of these programs, hampers the incorporation of these roles as part of reimbursed, routine care available to select or all populations. Although several state and national bills have already approved funding for patient navigation programs,[37-39] incorporation into Medicare or Medicaid services of more

widespread patient navigation systems for vulnerable populations requires stronger evidence of its benefits and costs.

The NCI PNRP is unique in examining the outcomes of care in patient navigation for persons across 4 different types of cancer, and across multiple diverse clinical care sites and populations. The study will assess the ability of patient navigation to facilitate timely and quality care from the time of the initial cancer screening abnormality through the completion of initial cancer therapy. By developing a core training program, this program will develop curricula we anticipate will be useful for navigator programs throughout the country. By recording and linking patient navigation activities between the navigator and each patient, we will be able to conduct secondary analyses on the effectiveness of navigation as a function of work load and activities of the navigator, and provide critical information regarding the optimal caseload for a navigator.

The PNRP emphasizes the importance of beginning measurement of time in care at the point of abnormal screening. To encompass all potential delays in care, we have defined our endpoints as time until definitive diagnosis and time to the initiation and completion of initial therapies. Our study will not have the power to assess changes in the stage of diagnosis or survival benefits of navigation. Benefits of navigation will be concluded from improvements in timeliness of care and completeness of treatment. Other studies have documented that the timeliness and completion of recommended therapy are associated with improvements in survival, especially in the elderly.[29,30,40] Our research study does not address the issues of screening, nor of survivorship after treatment.

The limitations of our methodology reflect the limitations inherent in research addressing dissemination of programs within community settings. The cooperative group includes both randomized clinical trials, which assign subjects to the intervention and control groups, and quasiexperimental designs, with assignment based on site of care. These differences reflect community and local needs when conducting community-based participatory research. Each methodology has its strengths and weaknesses in addressing the questions of interest in the research project. The randomized trial methodologies benefit from balance of known and unknown confounders between the 2 groups studied, but is limited in the generalizability to those subjects able to be reached and willing to be randomized. Those sites that include all subjects based on site of care risk confounding by site of care; however, by designing the intervention as a new standard of care that allows for the collection of data on all eligible subjects, they benefit in generalizability by the inclusion of those very subjects most difficult to reach and for whom the navigation intervention is designed to provide support. A second major limitation is the lack of power to address stage at diagnosis and survival outcomes and the need to use intermediate outcomes of the timeliness of completion of care and patient satisfaction.

The multidisciplinary approach to cancer care has resulted in significant survival gains, but at the cost of increased complexity within the healthcare system. The persistent gap in translating these improvements in cancer care to vulnerable populations will result in persistent and even widening racial disparities in cancer outcomes unless we develop and disseminate specific interventions to facilitate the process of care. Patient navigation represents a novel approach to addressing the barriers to the completion of cancer care in groups of patients vulnerable to inadequate care by virtue of their economic, cultural, educational, racial, and/or ethnic status.

REFERENCES

1. Brenner H, Gondos A, Arndt V. Recent major progress in long-term cancer patient survival disclosed by modeled period analysis. *J Clin Oncol*. 2007;25:3274-3280.
2. National Cancer Institute. Cancer Trends Progress Report — 2005 Update. Bethesda, MD: National Cancer Institute; 2005. Available from URL: http://progressreport.cancer.gov/ Accessed November 5, 2007.
3. Weir HK, 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.
4. Shavers VL, Fagan P, McDonald P. Health disparities across the cancer continuum. *J Health Care Poor Underserved*. 2007;18(4 suppl):1-5.
5. Peterson NB, Han J, Freund KM. Inadequate follow-up for abnormal Pap smears in an urban population. *J Natl Med Assoc*. 2003;95:825-832.
6. Maly RC, Umezawa Y, Ratliff CT, Leake B. Racial/ethnic group differences in treatment decision-making and treatment received among older breast carcinoma patients. *Cancer*. 2006;106:957-965.
7. Institute of Medicine. Unequal Treatment: Confronting Racial and Ethnic Disparities in Healthcare. Washington, DC: Institute of Medicine; 2003.
8. Freeman HP. Patient navigation: a community centered approach to reducing cancer mortality. *J Cancer Educ*. 2006;21(1 suppl):S11-S14.
9. Dohan D, Schrag D. Using navigators to improve care of underserved patients: current practices and approaches. *Cancer*. 2005;104:848-855.

10. 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):359-367.
11. 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.
12. Norris SL, Nichols PJ, Caspersen CJ, et al. The effectiveness of disease and case management for people with diabetes. A systematic review. *Am J Prev Med*. 2002;22(4 suppl):15-38.
13. National Cancer Institute. Patient Navigation Research Program RFA-CA005-019. Bethesda, MD: National Caner Institute; 2004. Available at: http://grants.nih.gov/grants/guide/rfa-files/RFA-CA-05-019.html Accessed January 3, 2008.
14. Longest B, Young GJ. Coordination and communication. In: Shortell S, Kaluzny AD, eds. Health Care Management: Organization Design and Behavior. Albany, NY: Thomson Delmar Learning; 2000:210-243.
15. Cross T, Bazron B, Dennis K, Isaacs M. Towards a Culturally Competent System of Care: A Monograph on Effective Services for Minority Children Who are Severely Emotionally Disturbed. Vol 1. Washington, DC: Georgetown University Child Development Center; 1989.
16. Calhoun E, Whitley EM, Esparza A, et al. A National Patient Navigator Training Program. Health Promotion and Practice (in press).
17. van der Aa MA, Schutter EM, Looijen-Salamon M, Martens JE, Siesling S. Differences in screening history, tumour characteristics and survival between women with screen-detected versus not screen-detected cervical cancer in the east of the Netherlands, 1992-2001. *Eur J Obstet Gynecol Reprod Biol*. 2008;139:204-209.
18. Abner AL, Collins L, Peiro G, et al. Correlation of tumor size and axillary lymph node involvement with prognosis in patients with T1 breast carcinoma. *Cancer*. 1998;83:2502-2508.
19. Andriole GL, Levin DL, Crawford ED, et al. Prostate Cancer Screening in the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial: findings from the initial screening round of a randomized trial. *J Natl Cancer Inst*. 2005;97:433-438.
20. NCCN Clinical Practice Guidelines in Oncology, Breast Cancer. National Comprehensive Cancer Network. Available at: http://www.nccn.org/professionals/physician_gls/PDF/breast.pdf Accessed January 3, 2008.
21. NCCN Clinical Practice Guidelines in Oncology, Cervical Cancer. National Comprehensive Cancer Network. Available at: http://www.nccn.org/professionals/physician_gls/PDF/cervical.pdf Accessed January 3, 2008.
22. NCCN Clinical Practice Guidelines in Oncology, Prostate Cancer. National Comprehensive Cancer Network. Available at: http://www.nccn.org/professionals/physician_gls/PDF/prostate.pdf Accessed January 3, 2008.
23. NCCN Clinical Practice Guidelines in Oncology, Colon Cancer. National Comprehensive Cancer Network. Available at: http://www.nccn.org/professionals/physician_gls/PDF/colon.pdf Accessed January 3, 2008.
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;197:346-355.
25. Benard VB, Lee NC, Piper M, Richardson L. Race-specific results of Papanicolaou testing and the rate of cervical neoplasia in the National Breast and Cervical Cancer Early Detection Program, 1991-1998 (United States). *Cancer Causes Control*. 2001;12:61-68.
26. 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;90:130-134.
27. Horowitz M, Wilner N, Alvarez W. Impact of Event Scale: a measure of subjective stress. *Psychosom Med*. 1979;41:209-218.
28. Wolf MS, Chang CH, Davis T, Makoul G. Development and validation of the Communication and Attitudinal Self-Efficacy scale for cancer (CASE-cancer). *Patient Educ Couns*. 2005;57:333-341.
29. Geiger AM, Thwin SS, Lash TL, et al. Recurrences and second primary breast cancers in older women with initial early-stage disease. *Cancer*. 2007;109:966-974.
30. Lash TL, Silliman RA, Guadagnoli E, Mor V. The effect of less than definitive care on breast carcinoma recurrence and mortality. *Cancer*. 2000;89:1739-1747.
31. Barabasi AL. Linked. New York: Plume; 2003.
32. Keating NL, Ayanian JZ, Cleary PD, Marsend PV. T Factors affecting influential discussions among physicians: a social network analysis of a primary care practice. *J Gen Intern Med*. 2007;22:794-798.
33. Norros L, Nuutinen M. The concept of the core task and the analysis of working practices. In: Boreham N, ed. Work Process Knowledge. New York: Routledge; 2002.
34. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis*. 1987;40:373-383.
35. Davis TC, Long SW, Jackson RH, et al. Rapid estimate of adult literacy in medicine: a shortened screening instrument. *Fam Med*. 1993;25:391-395.
36. Morris NS, MacLean CD, Chew LD, Littenberg B. The Single Item Literacy Screener: evaluation of a brief instrument to identify limited reading ability. *BMC Fam Pract*. 2006;7:21.
37. House Committee Report. Patient Navigation Outreach and Chronic Disease Prevention Act of 2005. 109th Cong, 1st Sess (2005).
38. House Committee Report. Patient Navigation Outreach and Chronic Disease Prevention Act of 2003. 108th Cong, 1st Sess (2003).
39. Bill 4705 Relative to Patient Navigators. Massachusetts House, Committee on Public Health. February 28, 2006. Available at: http://www.mass.gov/legis/184history/h04705.htm Accessed January 3, 2008.
40. Early Breast Cancer Trialists' Collaborative Group (EBCTCG). Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials. *Lancet*. 2005;365:1687-1717.

The Relative Efficacy of Meperidine for the Treatment of Acute Migraine: A Meta-analysis of Randomized Controlled Trials

Benjamin W. Friedman, MD, MSc
Alok Kapoor, MD, MSc
Matt S. Friedman, MD
Michael L. Hochberg, MD
Brian H. Rowe, MD, MSc

From the Department of Emergency Medicine, Albert Einstein College of Medicine, Bronx, NY (B. W. Friedman, Hochberg); the Hospital Medicine Unit, Section of General Internal Medicine, Boston University, Boston, MA (Kapoor); the Department of Emergency Medicine, Mount Sinai Medical Center, New York, NY (M. S. Friedman); and the Department of Emergency Medicine and Department of Public Health Sciences, School of Public Health, University of Alberta, Edmonton, Alberta, Canada (Rowe).

Study objective: Despite guidelines recommending against opioids as first-line treatment for acute migraine, meperidine is the agent used most commonly in North American emergency departments. Clinical trials performed to date have been small and have not arrived at consistent conclusions about the efficacy of meperidine. We performed a systematic review and meta-analysis to determine the relative efficacy and adverse effect profile of opioids compared with nonopioid active comparators for the treatment of acute migraine.

Methods: We searched multiple sources (Cochrane Central Register of Controlled Trials, MEDLINE, EMBASE, CINAHL, and LILACS, emergency and headache medicine conference proceedings) for randomized controlled trials comparing parenteral opioid and nonopioid active comparators for the treatment of acute migraine headache. Our primary outcome was relief of headache. If this was unavailable, we accepted rescue medication use or we transformed visual analog scale change scores by using an established procedure. We grouped studies by comparator: a regimen containing dihydroergotamine, antiemetic alone, or ketorolac. For each study, we calculated an odds ratio (OR) of headache relief and then assessed clinical and statistical heterogeneity for the group of studies. We then pooled the ORs of headache relief with a random-effects model.

Results: From 899 citations, 19 clinical trials were identified, of which 11 were appropriate and had available data. Four trials involving 254 patients compared meperidine to dihydroergotamine, 4 trials involving 248 patients compared meperidine to an antiemetic, and 3 trials involving 123 patients compared meperidine to ketorolac. Meperidine was less effective than dihydroergotamine at providing headache relief (OR=0.30; 95% confidence interval [CI] 0.09 to 0.97) and trended toward less efficacy than the antiemetics (OR=0.46; 95% CI 0.19 to 1.11); however, the efficacy of meperidine was similar to that of ketorolac (OR=1.75; 95% CI 0.84 to 3.61). Compared to dihydroergotamine, meperidine caused more sedation (OR=3.52; 95% CI 0.87 to 14.19) and dizziness (OR=8.67; 95% CI 2.66 to 28.23). Compared to the antiemetics, meperidine caused less akathisia (OR=0.10; 95% CI 0.02 to 0.57). Meperidine and ketorolac use resulted in similar rates of gastrointestinal adverse effects (OR=1.27; 95% CI 0.31 to 5.15) and sedation (OR=1.70; 95% CI 0.23 to 12.72).

Conclusion: Clinicians should consider alternatives to meperidine when treating acute migraine with injectable agents. [Ann Emerg Med. 2008;52:705-713.]

0196-0644/$-see front matter

doi:10.1016/j.annemergmed.2008.05.036

INTRODUCTION

Migraine headache is a disorder that afflicts millions of North American patients,[1] resulting in nearly 1 million presentations to emergency departments (EDs) annually.[2] Parenteral opioids are used in 51% of all migraine visits to US EDs and are the therapeutic agent used most commonly in Canadian EDs.[2,3] Meperidine, the individual opioid agent used most frequently, is administered in 36% of all US migraine visits.[2]

Several consensus guideline statements caution against the use of opioids as first-line treatment for migraine, citing lack of efficacy, adverse medication effects, the potential for addiction or ED recidivism, and development of medication overuse headache.[3-5] Supporting these consensus statements are multiple

Editor's Capsule Summary

What is already known on this topic

Despite consensus guidelines recommending alternatives, parenteral opioids and particularly meperidine are frequently used in the emergency department setting to treat acute migraine.

What question this study addressed

This systematic review of 11 eligible studies addresses whether nonopioid medications provide better headache relief with fewer adverse effects than meperidine.

What this study adds to our knowledge

This systematic review found that meperidine is less effective than antiemetics or dihydroergotamine and is associated with more sedation and dizziness; however, meperidine may provide more pain relief than ketorolac.

How this might change clinical practice

Antiemetics or dihydroergotamine is more effective than meperidine in the treatment of acute migraine.

small studies that often fail to demonstrate a clinically important difference between the opioid and the active comparator. Two previous systematic reviews of parenteral treatment for acute migraine combined opioids with other comparators, thus not addressing the independent efficacy of opioids.[6,7] Because opioids are frequently used for patients with acute migraine in the context of emergency care, we conducted a systematic review and meta-analysis to identify and compare the efficacy of opioids with nonopioid active comparators for the management of acute migraine headaches.

MATERIALS AND METHODS

Study Design

This was a systematic review and meta-analysis to determine the efficacy, adverse event profile, and frequency of recurrent headache after treatment with injectable opioids compared with other active agents for the treatment of acute migraine.

Our search strategy is summarized in Figure 1. Using this strategy, we searched the Cochrane Registry of Controlled Trials, MEDLINE, EMBASE, LILACS, and CINAHL from earliest indexing until April 2007. We identified unpublished research by searching through electronically published abstracts from national meetings of emergency medicine, neurology, and headache medicine societies from 1985 to 2007. We reviewed all references from identified trials, guideline statements, and on-topic reviews and consulted with experts in the fields of emergency and headache medicine. In addition, we searched for similar systematic reviews and meta-analyses and used the PubMed "related articles" feature for all identified trials. The MEDLINE search was updated in April 2008, and no additional studies meeting the inclusion criteria were identified.

1. migraine.mp.†
2. exp migraine/
3. headache$.mp.
4. 1 or 2 or 3
5. exp opioid analgesics/
6. meperidine/ or pethidine/ or (meperidine or pethidine).mp.
7. hydromorphone/ or hydromorphone.mp.
8. fentanyl/ or fentanyl.mp.
9. butorphanol/ or butorphanol.mp.
10. 5 or 6 or 7 or 8 or 9
11. 4 and 10
12. limit 11 to "therapeutics (sensitivity 2 or more terms)"

Figure 1. OVID MEDLINE search strategy used to identify citations for this review of opioids in acute migraine.*

*OVID Medline searches the same MEDLINE database as PubMed. Variations in suffixes used to access individual field such as subject headings and textwords are mentioned below. Strategies may be combined by connecting the line number with "and" or "or" to generate the intersection or union of the retrieved citations. We followed a similar approach when we searched the EMBASE (not shown).

†Slashes indicate subject heading (MeSH headings). Dollar sign implies any words or phrases containing characters to the left of it will be included. ".mp." Refers to the field keyword that recruits articles that contain the text to the left of it in the title, subject heading, or abstract. "exp" Explodes terms to include records linked to this term within a medical subject heading.

Studies were selected for inclusion if they were randomized controlled trials of an injectable opioid versus an active comparator for the treatment of acute migraine regardless of language of publication. "Injectable" was defined as administration through intravenous, intramuscular, or subcutaneous routes. "Acute migraine" was defined with criteria established by the International Classification of Headache Disorders.[8] If International Classification of Headache Disorders criteria were not applied or if the study predated these criteria, the study was included if a reasonable attempt had been made to include migraine headaches rather than all benign headaches. Studies were only included if they presented data on headache intensity within 2 hours of treatment.

Data Collection and Processing

One author (B.W.F.) screened all abstracts identified by the search for potential eligibility. If eligibility was possible, the article was requested and submitted to 2 other authors for review (M.S.F., M.L.H.). Primary data abstraction was performed by 2 authors (M.S.F., M.L.H.). Disagreements were resolved by consensus when possible or by review of a third author (B.W.F.).

Outcome Measures

The primary outcome for this analysis was relief of headache within 1 hour of medication administration. The original authors' definition of relief was used. If the rate of relief was not reported, we included use of rescue medication instead. If neither outcome was available, we transformed change in visual analog scale into a dichotomous outcome as discussed below. As secondary outcomes, we reported functional disability after medication administration, recurrence of the headache after initial treatment, and adverse effects associated with the study medication.

The Jadad score was calculated for each study, with standard methodology.[9] Two reviewers independently recorded the Jadad score. Disagreements were resolved by consensus.

Primary Data Analysis

In the primary analysis, we calculated the odds ratio (OR) and 95% confidence intervals (CIs) for headache relief in the opioid group versus the comparator for each included study. For studies reporting only mean change in visual analog scale and SD, we computed the mean difference per trial and pooled these by using standardized mean difference. The standard error of the standardized mean difference was calculated and then converted to an OR with 95% CI by using an established methodology.[10] If SD was not documented, we requested this information from study authors. If the corresponding author did not have this information, we imputed the value and tested the assumption in sensitivity analysis consistent with recommendations from *The Cochrane Handbook for Systemic Reviews of Interventions*.[10] We also calculated relative risk (RR) for each study and calculated a pooled RR for each comparison type. Because there is no procedure available to convert continuous outcomes to the RR directly, in certain cases, we approximated the RR from the OR by using a described method.[11] We emphasize that this is an approximation because the method relies on knowing the rate of headache relief in the comparator population, which we imputed from the remaining studies of that comparator type.

We grouped studies into the following categories for analysis, according to the class of the comparator drug: dihydroergotamine, alone or as part of a 2-drug regimen; antiemetic alone; or ketorolac. Within the groups, we determined the clinical and statistical heterogeneity of effect estimates. We made a conservative choice to pool studies with a random-effects analysis, although we also examined the fixed-effects analysis as a sensitivity analysis. We performed all analyses using RevMan version 4.3 (The Cochrane Collaboration, Copenhagen, Denmark).

RESULTS

Our search identified 899 abstracts. We requested 23 articles for full text review and ultimately identified 11 studies involving 12 comparisons that met our entry criteria.[12-22] One study included an arm with meperidine and another with butorphanol.[11] Because all the other trials included meperidine or its analog, pethidine, we did not aggregate the butorphanol data with the other studies. A flow diagram of the process of study selection is presented in Figure 2.

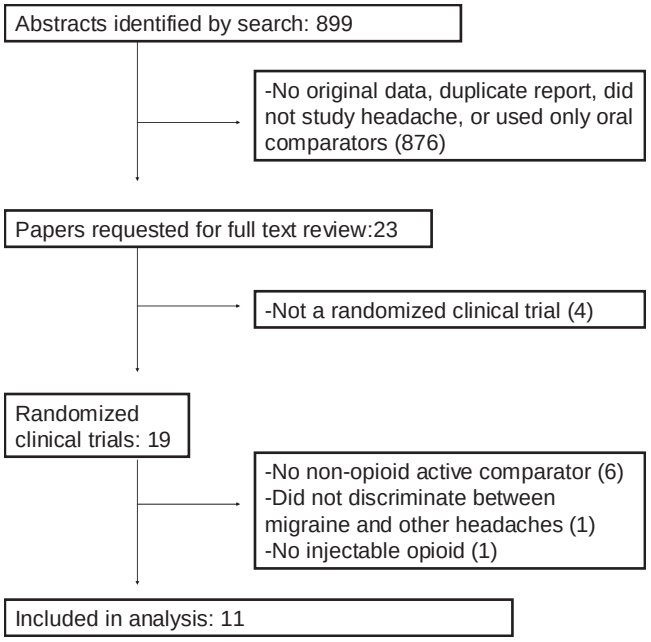


Figure 2. QUOROM flow diagram for selection of trials.[29]

Table 1 characterizes the included studies. Meperidine was used in doses ranging from 50 to 100 mg, though most commonly administered as 75 mg. In 2 studies, a weight-based dosing regimen was used. In 4 trials, meperidine was compared with a medication regimen containing dihydroergotamine, in 4 trials it was compared with an antiemetic alone, and in 3 trials it was compared with ketorolac. In 8 of the studies, meperidine was coadministered with an antihistamine per common clinical practice.

Four studies involving 254 patients compared meperidine to dihydroergotamine regimens. Two studies reported our primary outcome, proportion with headache relief.[13,17] One study reported mean difference in visual analog scale,[12] and the final study reported only the mean percentage improvement on a 5-point headache relief Likert scale.[21] The OR in each of these studies favored the dihydroergotamine regimen over meperidine, although this result was statistically significant in only 2 of the 4 studies. One study reported a sizable benefit of the dihydroergotamine regimen for headache relief (OR=0.02; 95% CI 0.0 to 0.23).[17] We could not find any distinguishing characteristic of that trial to explain the finding. One study used a higher dose of meperidine (1.5 mg/kg) and a dihydroergotamine regimen that did not contain metoclopramide[13]; these differences in study design likely explain why a larger benefit to dihydroergotamine was not found in that trial. The random-effects pooled estimate favored dihydroergotamine for headache relief (OR=0.30; 95% CI 0.09 to 0.97); however, the statistical heterogeneity was high

Table 1a. Characteristics of included studies.

First Author, y	Headache Type, Setting	N*	Opioid Dose	Administered With Opioid
Cicek, 2004	Vascular headache, ED	99	Pethidine 50 mg IM	None
Richman, 2002	IHS migraine, ED	29	Meperidine 1.5 mg/kg IM	None
Carleton, 1998	Vascular headache, ED	156	Meperidine 1.5 mg/kg IM	Hydroxyzine 0.7 mg/kg IM
Davis, 1995	Emergency physician–diagnosed migraine, ED	42	Meperidine 75 mg IM	Promethazine 25 mg IM
Scherl, 1995	Migraine criteria, clinic	27	Meperidine 75 mg IM	Promethazine IM
Klapper, 1993	Physician-diagnosed migraine, clinic	28	Meperidine 75 mg IM	Hydroxyzine 75 mg IM
Larkin, 1992	Emergency physician–diagnosed migraine, ED	31	Meperidine 75 mg IM	None
Duarte, 1992	IHS migraine, ED	50	Meperidine 100 mg IM	Hydroxyzine 50 mg IM
Stiell, 1991	Migraine score, ED	74	Meperidine 75 mg IM	Dimenhydrinate 50 mg IM
Belgrade, 1989	Migraine score, ED	43†	1) Meperidine 75 mg IM 2) Butorphanol 2 mg IM	1) Hydroxyzine 50 mg IM 2) None
Lane, 1989	Emergency physician–diagnosed migraine, ED	46	Meperidine 0.4 mg/kg×3 doses	Dimenhydrinate 25 mg

IM, Intramuscular; *IV,* intravenous; *Rescue medication,* requirement of rescue medication for persistent headache; *IHS,* International Headache Society's International Classification of Headache Disorders.
*N=total number of patients included in the primary efficacy analysis.
†Jadad score is a 5-point scale that measures methodological quality by assessing features such as method of randomization and allocation concealment. Scores above two are generally considered to represent higher quality studies.
‡In the Belgrade study, 22 patients received meperidine, 21 received DHE, and 19 received butorphanol.

(I^2=73%). The random-effects pooled RR was 0.53 (95% CI 0.24 to 1.17), meaning that meperidine provided headache relief about half as often as the dihydroergotamine regimens.

Four trials involving 248 patients compared 4 antiemetics to 4 doses of meperidine. One study reported the proportion with headache relief[18] and 3 studies reported use of rescue medication.[14,20,22] In 3 of the studies, the OR favored the antiemetic, though the OR was statistically significant in only 1 study. The remaining study used methotrimeprazine as its active comparator,[22] an antiemetic that is not commonly used for the treatment of migraine.[2] The random-effects pooled estimate favored the antiemetics for headache relief (OR=0.46; 95% CI 0.19 to 1.11); statistical heterogeneity was considerable (I^2=51%). The random-effects pooled RR was 0.81 (95% CI 0.66 to 1.00), meaning that meperidine provided headache relief 81% as often as the antiemetics.

Three trials involving 123 patients compared meperidine to ketorolac. The OR favored meperidine in all these studies, though never statistically significantly. The OR for headache relief using both a random-effects and fixed-effects analysis was 1.75 (95% CI 0.84 to 3.61); there was no significant statistical heterogeneity (I^2=0). The random-effects pooled RR was 1.23 (95% CI 0.90 to 1.68).

As detailed in Table 1, there were many methodological differences among the studies, including doses of meperidine ranging from 50 mg through 1.5 mg/kg; a variety of antihistamines coadministered with meperidine, such as promethazine, hydroxyzine, and dimenhydrinate; a spectrum of inclusion criteria, from physician-diagnosed migraine through formal application of established International Headache Society criteria; and several different outcomes that were incorporated in our analysis. Formal multivariate analyses were limited by insufficient statistical power (once we control for comparator class), so we examined trends in our data to determine whether variables such as study quality, criteria for enrollment, medication coadministered with meperidine, dose of meperidine, and outcome influenced our results. As is not uncommon in meta-analysis, we found that studies that had a more rigorous methodology and larger sample size tended to have an OR closer to 1. Specifically, with regard to enrollment criteria, the studies that used physician diagnosis of migraine rather than a "vascular" or migraine score had the most extreme values. Similarly, studies that reported details of their blinding and assignment methodology tended to have ORs closer to 1, though this was not universally true; Lane et al,[18] for example, who reported methodology in detail, had one of the most extreme OR outcomes. We could not identify any trend in outcome according to whether or not promethazine, hydroxyzine, or dimenhydrinate was coadministered with meperidine. For example, Richman et al,[20] who did not coadminister any medication with the meperidine, reported a relatively modest OR of 0.66. Lane et al,[18] who coadministered dimenhydrinate with the meperidine and Klapper et al,[17] who coadministered hydroxyzine, had more extreme ORs favoring the comparator. Similarly, a review of individual study ORs did not demonstrate a consistent outcome pattern according to dose of meperidine. Finally, we incorporated a variety of outcomes in our analysis. As we discuss in the "Limitations" section, this may have contributed to heterogeneity, but we believe this served to bias our analysis conservatively.

Functional disability was reported in 2 trials. Carleton et al[13] reported 14% of 75 patients randomized to meperidine and 32% of 74 patients randomized to dihydroergotamine were initially functionally impaired but then able to perform their usual activities 60 minutes after treatment (difference=18%; 95% CI 5% to 31%). Larkin et al[19] reported that 25% of 16 patients randomized to meperidine and 0 of 15 randomized to ketorolac could return to work unimpaired 60 minutes after treatment (difference=25%; 95% CI 3% to 47%).

Table 1b. Characteristics of included studies.

Comparator	Administered With Comparator	Time to Primary Outcome	Jadad Score†	Primary Outcome Used
Metoclopramide 10 mg IV	None	60 min	4	Rescue medication
Droperidol 2.5 mg IM	None	30 min	2	Rescue medication
DHE 1 mg IM	Hydroxyzine 0.7 mg/kg IM	60 min	5	Headache relief
Ketorolac 60 mg IM	None	60 min	5	Headache relief
DHE 0.5 mg IV	Metoclopramide	60 min	3	SMD change pain score
DHE 1 mg IV	Metoclopramide 10 mg IV	60 min	2	Headache relief
Ketorolac 30 mg IM	None	60 min	5	Headache relief
Ketorolac 60 mg IM	None	60 min	4	Headache relief
Methotrimeprazine 37.5 mg IM	None	60 min	5	Rescue medication
DHE 1 mg IV	Metoclopramide 1 mg IV	30 min	2	SMD visual analog scale change
Chlorpromazine 0.1 mg/kg	None	60 min	3	Headache relief

DHE, Dihydroergotamine; *SMD,* standardized mean difference.
*N=total number of patients included in the primary efficacy analysis.
†Jadad score is a 5-point scale that measures methodological quality by assessing features such as method of randomization and allocation concealment. Scores above two are generally considered to represent higher quality studies.
‡In the Belgrade study, 22 patients received meperidine, 21 received DHE, and 19 received butorphanol.

Recurrence of headache after treatment outcomes were reported in 3 studies. Carleton et al[13] reported 71% of 65 patients randomized to meperidine and 64% of 72 patients randomized to dihydroergotamine experienced headache during the 24 hours after ED discharge (difference=7%; 95% CI –9% to 23%). Scherl et al[21] reported 33% of 12 patients randomized to meperidine and 54% of 13 patients randomized to dihydroergotamine used additional pain medication within 24 hours (difference 21%; 95% CI –17% to 59%). Stiell et al[22] reported that 30% of 30 patients randomized to meperidine and 10% of 29 patients randomized to the antiemetic returned to the hospital for further treatment of headache (difference =20%; 95% CI 0, 40%).

A variety of adverse effects was reported for each comparison and is listed in Table 2. Compared with dihydroergotamine, meperidine caused more sedation (OR=3.52; 95% CI 0.87 to 14.19) and dizziness (OR=8.67; 95% CI 2.66 to 28.23). Compared with an antiemetic, meperidine caused less akathisia (OR=0.10; 95% CI 0.02 to 0.57).

No study reported long-term follow-up of the patients.

Parenteral butorphanol was compared with dihydroergotamine+metoclopramide in 1 trial.[12] There were no statistically significant or clinically relevant differences in efficacy or adverse effects between these 2 comparators.

LIMITATIONS

Some limitations of this review should be acknowledged. First, despite appropriate and similar patient eligibility criteria across included studies, it is likely that individuals with nonmigraine headache were enrolled in the trials. We attempted to restrict the included studies by using International Classification of Headache Disorders criteria; however, this was not always possible. We consider this a conservative bias that may have underestimated the treatment effect of the comparators, given the general efficacy of meperidine as an analgesic and the uncertain efficacy of dihydroergotamine and some antiemetics in nonmigraine headache syndromes.

Second, we retrieved few studies for each of the 3 comparisons. Small numbers prohibit robust conclusions. We could not therefore explore the effect of study level predictors such as dose of meperidine or coadministered antihistamines on pooled results. There was a fair amount of diversity with regard to these 2 study characteristics reflecting the range of clinical practice. Also, there is a possibility that our results are confounded by other external variables such as inclusion criteria, year of publication, methodologic quality, or outcome reported. We did notice a tendency toward more extreme ORs in less methodologically rigorous studies. Similarly, small numbers prohibit exploring the role of publication bias. Recent evidence suggests that publication bias is less pervasive in the ED literature[23]; however, negative trial results are less likely to be published and more likely to be excluded from a review of this nature, potentially biasing the study conclusions. We believe that our comprehensive search strategy, which included a hand search of recent conference proceedings to identify unpublished trials, minimized any such bias.

Third, selection and retrieval bias are always a concern in systematic reviews. However, all potentially relevant articles were screened by 2 independent reviewers with standardized eligibility criteria, decreasing the likelihood of this bias.

Finally, there was a substantial amount of clinical diversity in all 3 of our analyses and statistical heterogeneity in 2 of these analyses. Other authors who have performed similar analyses chose not to calculate the summary effect estimate because of disparate outcomes reported and statistical heterogeneity.[6,7] We were reassured about the integrity of our conclusions in the dihydroergotamine and antiemetic analyses by the fact that all but 1 of the study point estimates lay on the same side of the Forest plots as the summary ORs (Figure 3). Although magnitude of effect in the individual studies ranged from

Table 2. Summary effect measures of side effects of meperidine versus comparator in the treatment of acute migraine headache.

Side effect	Number of Studies	n/N Meperidine	n/N Comparator	I^2, %	Pooled OR (95% CI)
Meperidine vs dihydroergotamine					
Any adverse effect	2	40/100	47/99	82	0.49 (0.09–2.64)
Gastrointestinal	3	10/120	20/120	64	0.47 (0.10–2.34)
Sedation	3	37/113	21/113	55	3.52 (0.87–14.19)
Dizziness	3	22/113	3/113	0	8.74 (2.67–28.64)
Meperidine vs antiemetic					
Any adverse effect	2	34/71	30/74	78	1.06 (0.22–4.99)
Sedation	4	52/157	54/161	36	0.90 (0.42–1.94)
Dizziness	3	37/143	28/146	91	1.23 (0.08–17.91)
Akathisia, restlessness, or anxiety	3	0/135	12/137	0	0.11 (0.02–0.63)
Meperidine vs ketorolac					
Any adverse effect	1	12/25	7/25	NA	2.37 (0.73–7.68)
Gastrointestinal	2	5/41	4/40	0	1.27 (0.31–5.17)
Sedation	2	9/41	5/40	59	1.84 (0.88–3.84)

NA, Not applicable.
n=Number of patients reporting adverse effects summed across studies; N=at-risk population summed across studies.

clinically irrelevant to substantial, it is clear that meperidine is consistently less efficacious than these comparators. Unlike the previous work, we analyzed the proportion of patients requiring rescue medication when headache relief was not available. In the 4 trials that reported this outcome, we believe this to be an adequate surrogate for headache relief. Also, in the dihydroergotamine analysis, we transformed change in visual analog scale to an OR by using a standard procedure to enable meta-analysis with the other trials. As the transformed values decreased closer to the null when compared with studies in which the primary outcome was available, we believed that inclusion of these estimates was a conservative choice.

DISCUSSION

This systematic review and meta-analysis found that meperidine was significantly less efficacious than dihydroergotamine regimens for the treatment of acute migraine, caused more dizziness and sedation, and was less likely to result in return to normal functioning. There was a trend toward decreased efficacy of meperidine versus the antiemetics and a higher rate of return to the hospital in those who received meperidine, though the antiemetics caused a higher rate of akathisia. There were no substantial differences in efficacy or adverse event profile between meperidine and ketorolac.

Our conclusions are in keeping with guidelines published by the Canadian Association of Emergency Physicians and by the US Headache Consortium,[4,5] both of which present data from individual trials to argue that opioids should not be used as primary treatment for acute migraine. One systematic review, using many of the same clinical trials, concluded that dihydroergotamine administered with an antiemetic was as effective as, if not more effective than, multiple other treatments for acute migraine, including meperidine.[7] A systematic review of metoclopramide, an antiemetic, versus a host of comparators including meperidine concluded that patients who received metoclopramide were equally or more likely to report headache improvement.[6] Overall, the evidence here and elsewhere suggests that other agents are more effective than meperidine and produce fewer adverse effects.

Emergency medicine proponents of opioid use argue that these medications are highly effective, well tolerated, and safe and are important analgesics in a physician's armamentarium.[24-26] However, the data supporting these arguments primarily included patients with abdominal or flank pain and not headache patients. The wide variety of effective migraine-specific agents, such as the triptans, dihydroergotamine, and the antiemetics, among others, should relegate opioids to secondary status for the treatment of acute migraine. In contrast to the opioids, the migraine-specific agents seem to abort the pathologic process of the acute migraine, rather than just relieving the pain.[27]

Meperidine is commonly administered for the treatment of acute migraine in North American EDs. We have demonstrated that this practice is suboptimal. It remains unclear why emergency physicians use meperidine despite guideline statements to the contrary.[2] This may be related to patient-centered reasons such as familiarity with meperidine, a history of successful treatment of previous headaches with meperidine, or desire for an opioid-induced euphoria. Physicians may have concerns about adverse medication effects with alternate therapies, may be comfortable with dosing and administration of meperidine, or may not have been persuaded by a complicated evidence base. Clinicians should consider our results when choosing a therapy for an acute migraine, though they may also be guided by other factors such as contraindications and a patient's previous response to treatment. Some EDs have begun to eliminate meperidine from their formulary because of concerns about toxic metabolite accumulation and associated adverse events such as seizures. The effects of this action on treatment patterns in North American EDs have yet to be reported.

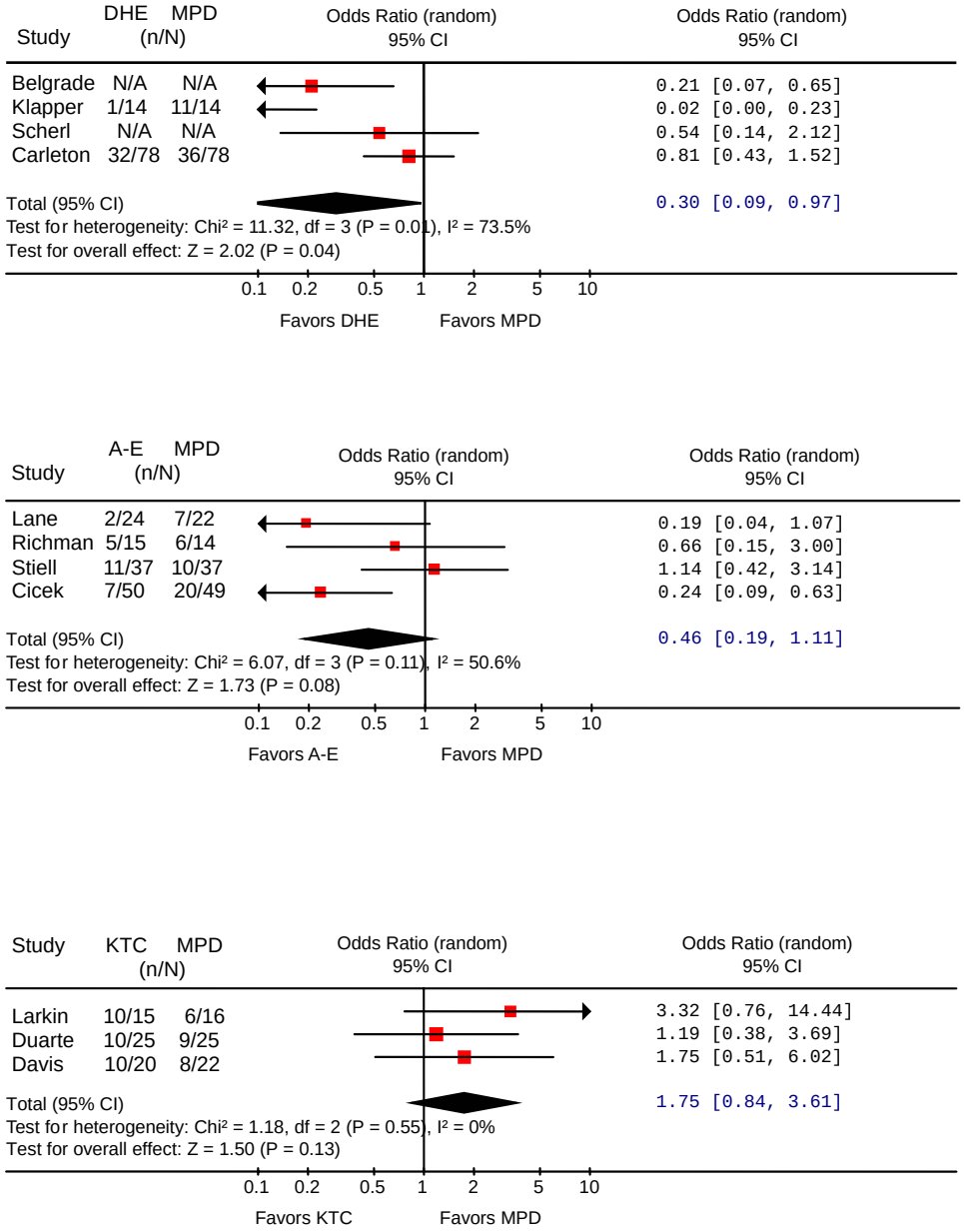


Figure 3. Forest plots of meperidine versus active comparators. These plots demonstrate odds of insufficient headache relief. *A,* A comparison of dihydroergotamine regimens versus meperidine N/A= not applicable, please see text. *B,* A comparison of anti-emetics versus meperidine. *C,* A comparison of ketorolac versus meperidine. *DHE,* Dihydroergotamine regimens; *MPD,* meperidine; *A-E,* anti-emetics; *KTC,* ketorolac.

Future research should measure provider choices and reasoning in different clinical scenarios and assess the effect of patient preferences.

Some issues remain unanswered. We did not find sufficient data to support or refute the contention that treatment of migraine with opioids was associated with medication overuse headache or addiction. Also, we did not identify any clinical trials using parenteral morphine or hydromorphone for the treatment of migraine or that compared sumatriptan to an opioid. Thus, the relative efficacy and adverse effect profile of these medications remain unknown. In the studies we identified, dihydroergotamine was often coadministered with metoclopramide, an agent with known efficacy in migraine.[6,28] It is unclear which agent was driving the efficacy and whether the 2 combined provided more headache relief than either alone. Similarly, it is unknown whether antihistamines coadministered with meperidine enhance efficacy. We did not have enough power to address this issue in our analysis; it is a hypothesis that will need to be answered in a clinical trial.

In conclusion, meperidine is less likely to relieve migraine than dihydroergotamine regimens and is associated with more adverse effects. There was also a trend toward decreased efficacy of meperidine compared with antiemetics. Insufficient evidence exists when comparing meperidine and ketorolac, though differences in efficacy between the 2 are minimal. Clinicians should consider alternatives to meperidine when treating acute migraine with injectable agents.

Supervising editors: Michael D. Brown, MD, MSc; Peter C. Wyer, MD

Author contributions: BWF and BHR identified the question of interest. AK and BWF designed the search strategy. BWF reviewed all relevant citations. BWF, MSF, and MLH abstracted data. BWF and AK reviewed and analyzed abstracted data. BWF, AK, and BHR interpreted and contextualized the results. BWF and AK drafted the article; all authors contributed substantially to editing and revision of the manuscript. BWF takes responsibility for the paper as a whole.

Funding and support: By *Annals* policy, all authors are required to disclose any and all commercial, financial, and other relationships in any way related to the subject of this article, that might create any potential conflict of interest. See the Manuscript Submission Agreement in this issue for examples of specific conflicts covered by this statement. Dr. Friedman is supported by a career development award (K23NS051049) from the National Institute of Neurological Disorders and Stroke. Dr. Rowe is supported by a 21st Century Research Chair Award from the government of Canada (Ottawa, Ontario).

Publication dates: Received for publication March 18, 2008. Revisions received April 22, 2008, and May 7, 2008. Accepted for publication May 20, 2008. Available online July 16, 2008.

Presented at the Society for Academic Emergency Medicine National Meeting, May 2008, Washington, DC.

Address for reprints: Benjamin W. Friedman, MD, MSc, Department of Emergency Medicine, Albert Einstein College of Medicine, Montefiore Medical Center, 111 East 210th Street, Bronx, NY 10467; 718-920-6626, fax 718-798-0730; E-mail befriedm@montefiore.org.

REFERENCES

1. Lipton RB, Scher AI, Kolodner K, et al. Migraine in the United States: epidemiology and patterns of health care use. *Neurology.* 2002;58:885-894.
2. Vinson DR. Treatment patterns of isolated benign headache in US emergency departments. *Ann Emerg Med.* 2002;39:215-222.
3. Colman I, Rothney A, Wright SC, et al. Use of narcotic analgesics in the emergency department treatment of migraine headache. *Neurology.* 2004;62:1695-1700.
4. Ducharme J. Canadian Association of Emergency Physicians Guidelines for the acute management of migraine headache. *J Emerg Med.* 1999;17:137-144.
5. Matchar DB, Young WB, Rosenberg JH, et al. Evidence-based guidelines for migraine headache in the primary care setting: pharmacological management of acute attacks. American Academy of Neurology practice guideline [online]. Available at: http://www.aan.com/professionals/practice/pdfs/gl0087.pdf. Accessed July 28, 2004.
6. Colman IBrown MD, Innes GD, et al. Parenteral metoclopramide for acute migraine: meta-analysis of randomised controlled trials. *BMJ.* 2004:329:1369-1373.
7. Colman IBrown MDInnes GD, et al. Parenteral dihydroergotamine for acute migraine headache: a systematic review of the literature. *Ann Emerg Med.* 2005:45:393-401.
8. Headache Classification Subcommittee of the International Headache Society. Classification and diagnostic criteria for headache disorders, cranial neuralgias and facial pain. Headache Classification Committee of the International Headache Society. *Cephalalgia.* 1988;8 (suppl 7):1-96.
9. Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials.* 1996:17:1-12.
10. Higgins J, Green S, eds. *The Cochrane Handbook for Systemic Reviews of Interventions, 4.2.6*. September 2006. Available at: http://www.cochrane.org/resources/handbook/Handbook4.2.6Sep2006.pdf. Accessed June 18, 2008. 2006; Chichester, UK: John Wiley and Sons, LTD.
11. Zhang J, Yu KF. What's the relative risk? a method of correcting the odds ratio in cohort studies of common outcomes. *JAMA.* 1998;280:1690-1691.
12. Belgrade MJ, Ling LJ, Schleevogt MB, et al. Comparison of single-dose meperidine, butorphanol, and dihydroergotamine in the treatment of vascular headache. *Neurology.* 1989;39:590-592.
13. Carleton SC, Shesser RF, Pietrzak MP, et al. Double-blind, multicenter trial to compare the efficacy of intramuscular dihydroergotamine plus hydroxyzine versus intramuscular meperidine plus hydroxyzine for the emergency department treatment of acute migraine headache. *Ann Emerg Med.* 1998; 32:129-138.
14. Cicek M, Karcioglu O, Parlak I, et al. Prospective, randomised, double blind, controlled comparison of metoclopramide and pethidine in the emergency treatment of acute primary vascular and tension type headache episodes. *Emerg Med J.* 2004;21: 323-326.
15. Davis CP, Torre PR, Williams C, et al. Ketorolac versus meperidine-plus-promethazine treatment of migraine headache: evaluations by patients. *Am J Emerg Med.* 1995;13:146-150.
16. Duarte C, Dunaway F, Turner L, et al. Ketorolac versus meperidine and hydroxyzine in the treatment of acute migraine headache: a randomized, prospective, double-blind trial. *Ann Emerg Med.* 1992;21:1116-1121.
17. Klapper JA, Stanton J. Current emergency treatment of severe migraine headaches. *Headache.* 1993;33:560-562.
18. Lane PL, McLellan BA, Baggoley CJ. Comparative efficacy of chlorpromazine and meperidine with dimenhydrinate in migraine headache. *Ann Emerg Med.* 1989;18:360-365.
19. Larkin GL, Prescott JE. A randomized, double-blind, comparative study of the efficacy of ketorolac tromethamine versus meperidine in the treatment of severe migraine. *Ann Emerg Med.* 1992;21:919-924.
20. Richman PB, Allegra J, Eskin B, et al. A randomized clinical trial to assess the efficacy of intramuscular droperidol for the treatment of acute migraine headache. *Am J Emerg Med.* 2002; 20:39-42.
21. Scherl ER, Wilson JF. Comparison of dihydroergotamine with metoclopramide versus meperidine with promethazine in the treatment of acute migraine. *Headache.* 1995;35:256-259.
22. Stiell IG, Dufour DG, Moher D, et al. Methotrimeprazine versus meperidine and dimenhydrinate in the treatment of severe migraine: a randomized, controlled trial. *Ann Emerg Med.* 1991; 20:1201-1205.

23. Ospina MB, Kelly K, Klassen TP, et al. Publication bias of randomized controlled trials in emergency medicine. *Acad Emerg Med.* 2006;13:102-108.
24. Birnbaum A, Esses D, Bijur PE, et al. Randomized double-blind placebo-controlled trial of two intravenous morphine dosages (0.10 mg/kg and 0.15 mg/kg) in emergency department patients with moderate to severe acute pain. *Ann Emerg Med.* 2007;49: 445-453, 453.e1-2.
25. Chang AK, Bijur PE, Meyer RH, et al. Safety and efficacy of hydromorphone as an analgesic alternative to morphine in acute pain: a randomized clinical trial. *Ann Emerg Med.* 2006: 48:164-172
26. Gallagher EJ, Esses D, Lee C, et al. Randomized clinical trial of morphine in acute abdominal pain. *Ann Emerg Med.* 2006;48: 150-160, 160.e1-4.
27. Goadsby PJ, Lipton RB, Ferrari MD. Migraine—current understanding and treatment. *N Engl J Med.* 2002;346:257-270.
28. Friedman BW, Corbo J, Lipton RB, et al. A trial of metoclopramide vs sumatriptan for the emergency department treatment of migraines. *Neurology.* 2005;64:463-468.
29. Moher D, Cook DJ, Eastwood S, et al. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet.* 1999:354:1896-1900.

Risk of Thromboembolism With Short-term Interruption of Warfarin Therapy

David A. Garcia, MD; Susan Regan, PhD; Lori E. Henault, MPH; Ashish Upadhyay, MD; Jaclyn Baker, MD; Mohamed Othman, MD; Elaine M. Hylek, MD, MPH

Background: Significant uncertainty surrounds the treatment of patients who must discontinue warfarin sodium therapy before an invasive procedure. In part, the uncertainty results from the lack of published information about the risk of thromboembolism associated with short-term warfarin therapy interruption. We aimed to assess the frequency of thromboembolism and bleeding within a large cohort of patients whose warfarin therapy was temporarily withheld for an outpatient invasive procedure.

Methods: This prospective, observational cohort study was performed at 101 sites (primarily community-based physician office practices) in the United States. Enrollment was conducted from April 4, 2000, to March 6, 2002. The main outcome measures were thromboembolism or clinically significant hemorrhage within 30 days of warfarin therapy interruption.

Results: A total of 1293 episodes of warfarin therapy interruption in 1024 individuals were included. The mean (SD) patient age was 71.9 (10.6) years; 438 (42.8%) were female. The most common indications for anticoagulant therapy were atrial fibrillation (n=550), venous thromboembolism (n=144), and mechanical heart valve (n=132). The most common procedures were colonoscopy and oral and ophthalmic surgery. Perioperative heparin or low-molecular-weight heparin was used in only 8.3% of cases overall. Seven patients (0.7%; 95% confidence interval [CI], 0.3%-1.4%) experienced postprocedure thromboembolism within 30 days. None of the 7 patients who experienced thromboembolism received periprocedural bridging therapy. Six patients (0.6%; 95% CI, 0.2%-1.3%) experienced major bleeding, whereas an additional 17 patients (1.7%; 95% CI, 1.0%-2.6%) experienced a clinically significant, nonmajor bleeding episode. Of these 23 patients who had bleeding episodes, 14 received periprocedural heparin or low-molecular-weight heparin. The duration of warfarin therapy interruption was variable; however, more than 80% of patients had warfarin therapy withheld for 5 days or fewer.

Conclusions: For many patients receiving long-term anticoagulation who need to undergo a minor outpatient intervention, a brief (≤5 days) periprocedural interruption of warfarin therapy is associated with a low risk of thromboembolism. The risk of clinically significant bleeding, even among outpatients undergoing minor procedures, should be weighed against the thromboembolic risk of an individual patient before the administration of bridging anticoagulant therapy.

Arch Intern Med. 2008;168(1):63-69

Author Affiliations: Department of Internal Medicine, University of New Mexico Health Sciences Center, Albuquerque (Drs Garcia and Othman); and Department of Medicine, General Medicine Unit, Massachusetts General Hospital (Dr Regan), Department of Medicine, Research Unit–Section of General Internal Medicine, Boston University School of Medicine (Drs Baker and Hylek and Ms Henault), and Department of Medicine, Caritas St Elizabeth's Medical Center, Tufts University School of Medicine (Dr Upadhyay), Boston.

HEALTH CARE PROFESSIONALS face a dilemma when a warfarin sodium–treated patient needs to undergo an elective procedure or minor surgery. In these circumstances, the risk of bleeding, if the procedure is performed without stopping warfarin therapy, must be weighed against the risk of thrombosis associated with warfarin therapy interruption. The patient and physician have 3 options: (1) continue warfarin therapy, (2) withhold warfarin therapy for some time before (and after) the procedure, or (3) temporarily withhold warfarin therapy while also providing a short-acting (bridging) anticoagulant (such as unfractionated heparin or low-molecular-weight heparin) during the perioperative period. Current guidelines from the American College of Chest Physicians suggest that if the annual risk of thromboembolism is low, warfarin therapy may be held for 4 to 5 days before the procedure and restarted shortly thereafter.[1] The 2006 guidelines for the treatment of patients with atrial fibrillation from the American College of Cardiology, American Heart Association, and European Society of Cardiology suggest an interval of up to 1 week without substituting heparin. The authors acknowledge that this level C recommendation is "based on extrapolation from the annual rate of thromboembolism"[2(p298)] and is not evidence based because no studies are available to inform this question.[2] For patients at higher risk for thromboembolism, several studies[3-6] have described outcomes in patients treated with periprocedural low-molecular-weight heparin. However, since none of these studies included a control arm (ie, patients for whom bridging therapy was not prescribed), the risk of thromboembolism associated with short-term warfarin therapy interruption

Table 1. Clinical Characteristics of Patients

Characteristic	Patients, No. (%) (n=1024)	Interruptions, No. (%) (n=1293)
Primary warfarin indication		
Atrial fibrillation	550 (53.7)	690 (53.4)
Venous thromboembolism	144 (14.1)	201 (15.5)
Diagnosed within 4 weeks	3 (2.1)	4 (1.9)
Prosthetic valve[a]	132 (12.9)	159 (12.3)
Mitral position	54 (40.9)	65 (40.9)
Aortic position	66 (50.0)	81 (50.9)
Unspecified position	12 (9.1)	13 (8.2)
Stroke[b]	93 (9.1)	117 (9.1)
Left ventricular dysfunction	34 (3.3)	43 (3.3)
Other	71 (6.9)	83 (6.4)
Active malignancy	16 (1.6)	22 (1.7)
Thromboembolic risk high[c] (interruption reason)	73 (7.1)	91 (7.0)
Colonoscopy	272 (26.6)	324 (25.1)
Oral or dental surgery	257 (25.1)	323 (24.9)
Ophthalmic surgery	89 (8.7)	116 (8.9)
Other[d]	406 (39.7)	530 (40.9)
Management of warfarin therapy[e]		
Preprocedure phytonadione	12 (1.2)	13 (1.0)
Bridging therapy with low-molecular-weight heparin	88 (8.6)[f]	108 (8.3)

[a] One patient had a prosthetic valve in both the mitral and aortic positions.
[b] Without documented atrial fibrillation or prosthetic heart valve.
[c] Includes prosthetic valve in the mitral position, venous thromboembolism within 4 weeks before warfarin interruption, and active malignancy.
[d] Common examples from this category included prostate or breast biopsy, epidural injection, and dermatologic procedures such as removal of skin cancer.
[e] The median length of withholding warfarin therapy was 3 days. Length of interruption was known for 886 patients (86.5%) and 1130 interruptions (87.4%).
[f] For patients who had more than 1 interruption, bridging therapy during the first interruption is presented here.

remains unknown. This represents a critical gap in current knowledge because the increased risk of hemorrhage associated with perioperative heparin is justified by the theory that such bridging therapy will prevent potentially devastating thromboembolic events (eg, stroke) that would otherwise occur. Without knowing the risk of thromboembolism associated with warfarin therapy interruption alone, an informed risk-benefit examination of bridging therapy cannot be performed. Three small observational studies[7-9] of patients with mechanical heart valves (n=28, n=16, and n=25) have reported successful warfarin therapy interruption without bridging therapy around the time of a procedure; however, the small size of these studies prohibits definitive conclusion.

The uncertainty surrounding periprocedural treatment strategies for patients who require warfarin therapy interruption has led to significant practice variation, even among experienced health care professionals who perform anticoagulation.[10,11] For example, in a 2003 survey[11] that involved hypothetical cases, 324 North American health care professionals were asked to select 1 of 4 perioperative anticoagulation strategy options. The overall lack of consensus found in the survey is highlighted by 1 case scenario in which each of the 4 possible strategies were chosen by 11%, 50%, 16%, and 22% of respondents, respectively.

In the present prospective observational study, we sought to quantify the risk of thromboembolism and bleeding associated with the short-term interruption of warfarin therapy for minor procedures. Using data gathered from primarily community-based physician practices, we also determined the number of days warfarin therapy was withheld and recorded the frequency with which low-molecular-weight heparin was used as a bridging anticoagulant for patients undergoing outpatient interventions.

METHODS

OVERVIEW

The overall objective of the Anticoagulation Consortium to Improve Outcomes Nationally (ACTION) study[12] was to assemble a large prospective cohort of patients taking warfarin to identify practice variations in warfarin therapy management that might lead to improved drug safety. Specific areas to be studied included frequency of monitoring, response to out-of-range international normalized ratio values, periprocedural warfarin therapy management, and overall quality of anticoagulation control and related outcomes across the United States. At the time this study was planned, a software program designed to aid in warfarin therapy management had already been in use in the United States for many years. This program, called CoumaCare (Bristol-Myers Squibb [formerly DuPont Pharmaceuticals], Princeton, New Jersey), was available at no cost, and technological support was also provided without charge. CoumaCare was designed to help physicians manage warfarin therapy and was used for clinical purposes to aid in patient tracking, data entry, and record keeping. The program did not make dosing or follow-up interval recommendations. The study protocol was approved by the Western Institutional Review Board, Olympia, Washington, and local review boards where they existed.

SITE RECRUITMENT

Established physician practices that used CoumaCare as the patient anticoagulation medical record were invited to participate through a study Web site (www.ACTIONregistry.com). The only material incentive for participation in the project was an informational quarterly newsletter written by the investigators. An independent registry specialist, McKesson HBOC BioServices (Rockville, Maryland), was responsible for all operational aspects of the study. McKesson is a health care services company that provides biomedical support services to the US government, pharmaceutical and biotechnology industries, universities, institutions, and contract research organizations.

A total of 174 individual site registrations were received by McKesson. Of these, 101 sites had the technological capability and the institutional review board approval necessary to participate and were enrolled. All sites had at least 1 dedicated professional managing warfarin therapy, usually within the setting of a community-based, physician group practice. McKesson provided individual on-site training related to all aspects of the research protocol, such as patient recruitment, informed consent procedures, data entry, and transmission. Adverse event reporting was mandatory, and study personnel were trained to perform such reporting with rigor sufficient to meet federal regu-

Table 2. Bridging Therapy by Indication

Primary Indication	No. of Patients	Received Bridging Therapy (on First Interruption), No. (%)	No. of Total Interruptions	Received Bridging Therapy, No. (%)
Atrial fibrillation	550	15 (2.7)	690	17 (2.5)
Venous thromboembolism	144	15 (10.4)	201	22 (10.9)
Prosthetic valve	132	38 (28.8)	159	44 (27.7)
Stroke	93	7 (7.5)	117	10 (8.6)
Left ventricular dysfunction	34	2 (5.9)	43	2 (4.7)
Other	71	11 (15.5)	83	13 (15.7)
Total	**1024**	**88** (8.6)	**1293**	**108** (8.4)

latory requirements. Patients were invited to participate by letter, by clinic flyer, or in person (at the time of a routine appointment). To be eligible, patients had to be 18 years or older and provide written informed consent. Enrollment was conducted from April 4, 2000, to March 6, 2002.

DATA MANAGEMENT

Encrypted data from each site were downloaded by modem and transmitted to the independent data coordinating center weekly. Missing data fields and questionable values were flagged and resolved directly with the sites before data were transferred to the study investigators. Study investigators were blinded to the identification and location of participating practices and patients.

PATIENTS AND OUTCOMES

Episodes of warfarin therapy interruption were identified by direct review of the 102 732 free-text patient treatment notes. For interruptions to be included, available documentation had to state that warfarin therapy was being interrupted because a procedure was planned, and subsequent documentation had to indicate that the planned procedure occurred (ie, was not canceled or postponed indefinitely). Surgical procedures (including cardiovascular interventions) that required hospitalization were not eligible because responsibility for the management of anticoagulation would no longer be under the purview of the clinic and a reliable determination about heparin exposure could not be made.

The primary outcomes of interest were thromboembolism or hemorrhage within the 30-day period after the date of warfarin therapy interruption. Major hemorrhage was defined as bleeding that was fatal, led to hospitalization with transfusion of at least 2 U of packed red blood cells, or occurred at a critical site (eg, intracranial or retroperitoneal). Clinically significant, nonmajor hemorrhage was defined as other bleeding that led to an unplanned medical intervention (eg, subsequent operation or nasal packing). All primary outcome events were abstracted by an investigator (D.A.G. or E.M.H.) from patient treatment notes and validated by supporting information obtained by McKesson. For each patient with an identified interruption before a procedure, the following data were also subsequently abstracted: type of procedure, number of days warfarin therapy was withheld, presence or absence of bridging therapy (eg, low-molecular-weight heparin or unfractionated heparin), and the presence or absence of phytonadione use for warfarin reversal. Patients whose final study international normalized ratio was transmitted to the data coordinating center 30 days or fewer after their warfarin therapy was held were excluded, since complete follow-up surveillance for bleeding or thromboembolic events could not be ensured.

STATISTICAL ANALYSIS

Exact 95% confidence intervals (CIs) were calculated for the proportion of patients who sustained a thromboembolic or hemorrhagic event using the Poisson distribution.

RESULTS

In total, 1584 episodes of warfarin therapy interruption were identified. Of these, 1293 were included in our analysis (we did not include 228 inpatient vascular procedures, 62 major surgical procedures, or the 1 patient who was lost to follow-up). For the 1024 patients among whom the 1293 included procedures were performed, the mean (SD) age was 71.9 (10.6) years; 438 (42.8%) were female. The most common indications for warfarin use among the included patients were atrial fibrillation (n=550), venous thromboembolic disease (n=144), and mechanical heart valve (n=132) (**Table 1**). Of the valves, 54 (40.9%) were in the mitral position, 66 (50.0%) were aortic, and 12 (9.1%) were unspecified. Overall, 73 patients (7.1%) would have been considered high risk for thromboembolism because they had 1 or more of the following: prosthetic valve in the mitral position, venous thromboembolism within 4 weeks, or active malignancy. The most common reasons to interrupt warfarin therapy were colonoscopy (n=324), oral or dental surgery (n=323), and ophthalmic surgery (n=116). Among the remaining 530 interruptions classified as being due to "other procedure/biopsy/minor surgery," the most common reasons for withholding warfarin therapy were epidural injection, prostate biopsy, breast biopsy, and dermatologic procedures.

PERIPROCEDURAL USE OF HEPARIN

Perioperative treatment with heparin, which was almost exclusively low-molecular-weight heparin, was documented for 88 of the 1024 patients undergoing their first elective interruption (8.6%) and for 108 of the 1293 total interruptions (8.3%) (**Table 2**). Among the 550 patients with atrial fibrillation, 15 (2.7%) received periprocedural low-molecular-weight heparin. For patients with prosthetic valves, 38 of 132 (28.8%) received bridging therapy with low-molecular-weight heparin.

The duration of warfarin therapy interruption was available for 1130 episodes (87.4%). The frequency with which different interruption intervals were chosen is shown in **Figure 1**. In 947 of the 1130 interruptions (83.8%), warfarin therapy was withheld for 5 days or fewer; warfarin therapy was interrupted for more than 7 days in 71 instances (6.3%). Preprocedural use of phytonadione for reversal of anticoagulation was documented in only 13 cases (1.0%).

THROMBOEMBOLISM AND HEMORRHAGE

Of the 1024 patients, 7 (0.7%; 95% CI, 0.3%-1.4%) sustained a thromboembolic event during the 30-day follow-up period; 4 were arterial and 3 were venous (**Table 3**). None of the 7 patients with thrombosis received periprocedural bridging therapy; 2 would have been considered high risk (recent venous thromboembolism and active malignancy). If the patients who received no perioperative heparin are considered separately, the proportion with a thromboembolic event is unchanged (7/936, 0.7%; 95% CI, 0.3%-1.5%). When all 1293 interruptions are considered, the proportion associated with thromboembolism within the 30-day follow-up period is 0.5% (7/1293; 95% CI, 0.3%-1.1%). Among patients whose warfarin therapy was interrupted for 5 days or fewer, the proportion experiencing thromboembolism was 0.4% (4/984; 95% CI, 0.2%-1.0%) compared with 2.2% (3/135; 95% CI, 0.8%-6.3%) for those with an interruption interval of 7 days or more.

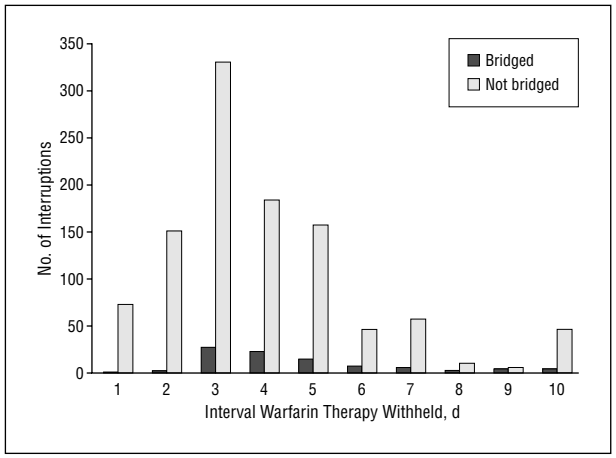


Figure 1. Overview of the frequency with which different interruption intervals were chosen.

After the procedure, 6 patients (0.6%; 95% CI, 0.2%-1.3%) sustained a major hemorrhage (1 spontaneous subdural, 4 gastrointestinal after colonoscopy, and 1 soft tissue with compartment syndrome), and 17 patients (1.7%; 95% CI, 1%-2.6%) experienced clinically significant, nonmajor bleeding. Four of the 6 patients with major bleeding episodes and 10 of the 17 patients with clinically significant, nonmajor bleeding received periprocedural heparin. **Figure 2**A presents the bleeding and thromboembolic events according to whether the patients received bridging therapy. Although the patients who received perioperative anticoagulation represented only 8.6% of the total population, 4 of the 6 major hemorrhagic events occurred in this group. Of note, none of the thromboembolic events occurred after a bleeding episode.

Among the 550 patients with atrial fibrillation, 4 (0.7%; 95% CI, 0.2%-1.9%) experienced an arterial thromboembolic event (3 strokes and 1 systemic embolism, probable ischemic bowel). One of the stroke events occurred 30 days after a 7-day warfarin therapy interruption. Of the 550 patients, 15 received heparin. If the patients with atrial fibrillation who received no perioperative bridging therapy are considered separately, the proportion of patients with an arterial thromboembolic event is unchanged (4/535, 0.7%). For each patient with atrial fibrillation, the number of risk factors for stroke[13] was determined. Use of transition therapy stratified by stroke risk factors is indicated in **Table 4**. Our data suggest that patients with a prior stroke may be at heightened risk, but the number of events is too small to draw any definitive conclusion.

Two patients sustained a major hemorrhage and 4 patients experienced clinically significant, nonmajor bleeding during the 30 days after their procedure. One of these 6 patients had received heparin therapy. Outcome data for this subgroup of patients with atrial fibrillation, stratified by whether bridging therapy was used, are displayed in Figure 2B.

Table 3. Clinical Details of Thromboembolic Events

Thromboembolic Event[a]	Warfarin Indication	Age, y/ Sex	Procedure	Days, No.		Comment	Phytonadione Reversal
				Warfarin Withheld	Event		
Minor stroke	AF	72/F	Shoulder procedure	4	6		No
Minor stroke	AF	75/F	Foot surgery	4	10		No
Ischemic bowel (probable)[b]	AF	73/F	Colonoscopy	5	11	Died within 30 days	No
Stroke	AF	82/F	Oral surgery	7	30		No
Pulmonary embolism	AF	72/M	EGD	5	9	Active cancer	No
DVT	DVT	53/F	Sinus surgery	10	15		No
Pulmonary embolism	DVT	69/F	Myelogram	10	10	Distal DVT 2-3 wk earlier	Yes

Abbreviations: AF, atrial fibrillation; DVT, deep venous thrombosis; EGD, esophagogastroduodenoscopy.
[a] None of these patients received periprocedural anticoagulation with heparin.
[b] Patient experienced the acute onset of abdominal pain 6 days after colonoscopy. The international normalized ratio at the time of presentation was 1.5. Clinical diagnosis was ischemic bowel, and the patient was treated conservatively. Clinical status subsequently worsened, and the patient died 4 days later (27 days after the procedure).

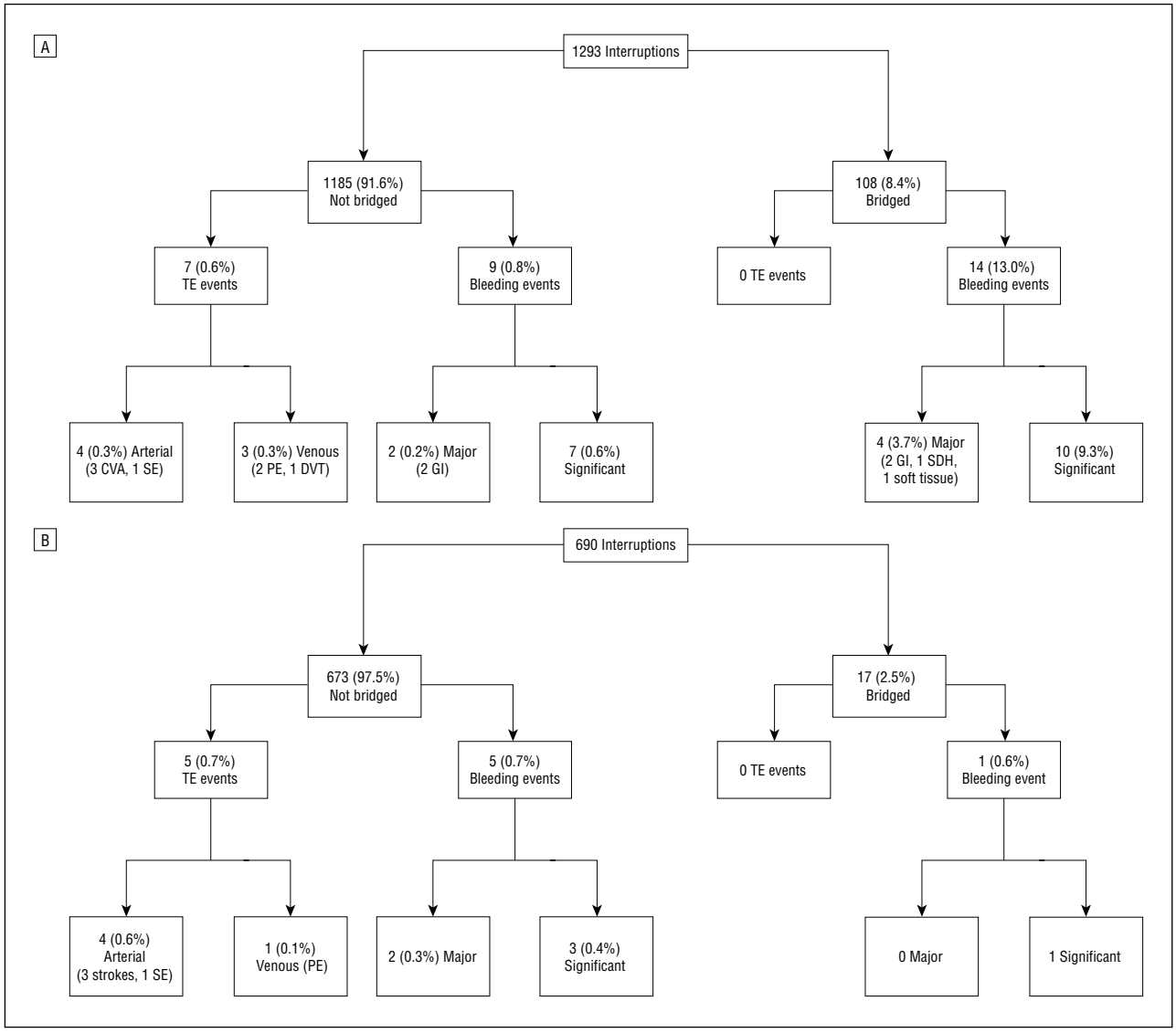


Figure 2. Thirty-day outcome events by bridging status for interruptions. A, Events for all patients. B, Events for patients with atrial fibrillation. CVA indicates cerebrovascular accident; DVT, deep venous thrombosis; GI, gastrointestinal; PE, pulmonary embolism; SDH, subdural hematoma; SE, systemic embolism; and TE, thromboembolism.

COMMENT

Among 6761 prospectively enrolled patients, we identified 1293 episodes of warfarin therapy interruption for elective outpatient procedures in 1024 patients. Of the 1293 interruptions, 108 (8.4%) were bridged with heparin or low-molecular-weight heparin. Most patients in our study would be considered to be at low to intermediate thromboembolic risk, which likely explains the infrequent use of heparin therapy. Patients at highest risk for thromboembolism composed only 7% of our cohort. The overall proportion of interruptions associated with thromboembolism within the 30-day follow-up period was 0.5% (7/1293). None of these patients had received heparin; 2 would have been considered high risk (recent venous thromboembolism and active malignancy). Among patients whose warfarin therapy was interrupted for 5 days or fewer, the proportion with throm-

Table 4. Proportion of Patients With Atrial Fibrillation Who Received Bridging Therapy According to Stroke Risk Factors

Stroke Risk Factor	No. of Patients (n=550)	Received Bridging Therapy, No. (%) (n=15)	Arterial Events, No. (%) (n=4)[a]
Age >75 y	285	4 (1.4)	2 (0.7)
Prior stroke	60	6 (9.8)	2 (3.3)
Hypertension	277	9 (3.2)	2 (0.7)
Diabetes mellitus	104	1 (0.96)	1 (0.96)
Heart failure	139	7 (5)	1 (0.7)
No. of risk factors			
0	75	2 (2.7)	0
1	191	5 (2.6)	0
2	194	4 (2.0)	4 (2.1)
3	75	2 (2.7)	0
≥4	15	2 (13.3)	0

[a] None of these patients received bridging therapy.

boembolism was 0.4% (4/984). No thromboembolic events occurred among the 108 interruptions bridged with heparin. Major hemorrhage was uncommon but was higher among patients receiving heparin (4/108, 3.7%) than among those patients who did not receive heparin (2/1185, 0.2%). The proportion of patients with clinically significant, nonmajor bleeding was also higher in those who received periprocedural heparin (9% vs 0.6%, respectively).

Our study demonstrates that thromboembolism is uncommon among low- to intermediate-risk outpatients who undergo elective periprocedural warfarin therapy interruption for a brief period (≤5 days). Our findings support a previous proposal that perioperative anticoagulation may be unnecessary for a significant proportion of patients who have undergone long-term anticoagulation whose warfarin therapy must be interrupted.[14] Our results are consistent with current guidelines, proposed by the American College of Chest Physicians, suggesting that low-risk patients may undergo 4 to 5 days of warfarin therapy interruption without bridging therapy.[1]

The evidence from our study (and others[4,5]) that bridging therapy may result in significant and potentially avoidable perioperative hemorrhage emphasizes the need for a randomized controlled trial of heparin vs placebo among warfarin-treated patients who need procedures. The principal challenge of such a trial will be to enroll a large number of participants while avoiding selection bias during the screening process.

Our study has several limitations. First, the observational nature of our study leaves open the possibility of selection bias, whereby only the patients at high risk for thromboembolism may have received heparin. However, because less than 10% of the overall cohort underwent transition with heparin, we do not believe that such a selection bias would substantively change our results. For example, among patients with atrial fibrillation, there were 690 interruptions; 17 (2.5%) were bridged. Even if one conservatively assumes that, in the absence of bridging therapy, 2 of these 17 patients (11.8%) would have experienced thrombosis, the overall proportion experiencing an arterial thromboembolism would increase from 0.6% (4/690) to 0.9% (6/690). In contrast, among patients with mechanical prosthetic heart valves, bridging therapy was administered in 28.8% of cases. Thus, we acknowledge that for patients with mechanical heart valves, our analysis may, because of selection bias, underestimate the risk of thromboembolism associated with warfarin therapy interruption in this subpopulation. Similarly, patients with a history of stroke constituted only 10.9% of our atrial fibrillation cohort. Because this subgroup is known to be at higher risk for stroke, further study of such patients is warranted.

We are confident that our study has captured all important events because our database was constructed by directly downloading patients' anticoagulation records. Because data were collected prospectively, the consecutive patients included in this study constitute an inception cohort established at the time warfarin therapy was discontinued. Since documentation is central to medical care (and of particular clinical importance in the hazard-prone area of anticoagulation), we highly doubt that health care professionals would neglect to enter information about prescribed warfarin therapy interruptions, major hemorrhage, thrombotic events, or the use of heparin. Of the 1024 patients identified in our study, only 1 patient was lost to follow-up.

It is possible that selection bias may have been introduced by voluntary participation and the requirement for written informed consent. However, the distribution of stroke risk factors in our overall cohort is similar to that seen in other published studies. The overall comparability provides reassurance that the patient population in our study is representative of patients with atrial fibrillation treated with warfarin.

Despite the unprecedented size of our study, the small number of thromboembolic events limits our ability to draw definitive conclusions about the risk of perioperative warfarin therapy interruption for any individual patient. The risk of perioperative thromboembolism is influenced by multiple patient-specific, procedure-specific, and physician-specific factors that collectively influence thrombogenicity. Quantifying the effect of independent factors (eg, indication for warfarin) with any degree of precision would require substantially more events. Our results, however, do not apply to patients undergoing major surgery or other more invasive procedures that require hospitalization. Compared with the procedures included in our study, more invasive surgical procedures would almost certainly be associated with a greater risk of bleeding. More invasive operations would also increase exposure of tissue factor to circulating plasma. Because this has the potential to increase coagulation activation[15] while reducing fibrinolytic capacity,[16] it is possible that such major surgical procedures would also confer a higher risk of thromboembolism.

We conclude that for many patients receiving long-term anticoagulation who need to undergo a minor outpatient intervention, a brief (≤5 days) periprocedural interruption of warfarin therapy is associated with a low risk of thromboembolism. Because our study and others[4] indicate that bridging anticoagulant therapy may be associated with a significantly increased risk of hemorrhage, a prospective randomized trial of bridging vs no bridging is needed to assess the risks and benefits of providing periprocedural anticoagulation. Until such a trial is completed, our results provide valuable information to physicians who must weigh the risks and benefits of different perioperative treatment strategies for patients taking warfarin.

Accepted for Publication: August 18, 2007.
Correspondence: David A. Garcia, MD, Department of Internal Medicine, 1 University of New Mexico, MSC10 5550, Albuquerque, NM 87131.
Author Contributions: Dr Hylek 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:* Garcia and Hylek. *Acquisition of data:* Garcia, Henault, Upadhyay, and Othman. *Analysis and interpretation of data:* Garcia, Regan, Henault, Upadhyay, Baker, Othman, and Hylek. *Drafting of the manuscript:* Garcia, Henault, Baker, Othman, and Hylek. *Critical revision of the manuscript for important intellec-*

tual content: Garcia, Regan, Upadhyay, and Hylek. *Statistical analysis:* Regan. *Obtained funding:* Hylek. *Administrative, technical, and material support:* Garcia, Henault, Baker, and Othman. *Study supervision:* Hylek.

Financial Disclosure: Dr Garcia has received consulting honoraria and research support from Bristol-Myers Squibb, AstraZeneca, and Sanofi-Aventis. Dr Hylek has served on advisory boards for Bristol-Myers Squibb and has received research support from Astra Zeneca and Bristol-Myers Squibb.

Funding/Support: This study was funded by Bristol-Myers Squibb (formerly Dupont Pharmaceuticals).

Role of the Sponsors: The sponsor had no role in any of the following: design and conduct of the study; the collection, management, analysis, and interpretation of the data; the preparation, review, or approval of the manuscript; or the decision to publish the findings.

Additional Contributions: Mark A. Crowther, MD, performed a helpful review of the manuscript.

REFERENCES

1. Ansell J, Hirsh J, Poller L, Bussey H, Jacobson A, Hylek E. The pharmacology and management of the vitamin K antagonists: the Seventh ACCP Conference on Antithrombotic and Thrombolytic Therapy [published correction appears in *Chest*. 2005;127(1):415-416]. *Chest*. 2004;126(3)(Suppl):204S-233S.
2. 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-e354.
3. Ferreira I, Dos L, Tornos P, Nicolau I, Permanyer-Miralda G, Soler-Soler J. Experience with enoxaparin in patients with mechanical heart valves who must withhold acenocumarol. *Heart*. 2003;89(5):527-530.
4. Kovacs MJ, Kearon C, Rodger M, et al. Single-arm study of bridging therapy with low-molecular-weight heparin for patients at risk of arterial embolism who require temporary interruption of warfarin. *Circulation*. 2004;110(12):1658-1663.
5. Douketis JD, Johnson JA, Turpie AG. Low-molecular-weight heparin as bridging anticoagulation during interruption of warfarin: assessment of a standardized periprocedural anticoagulation regimen. *Arch Intern Med*. 2004;164(12):1319-1326.
6. Spandorfer JM, Lynch S, Weitz HH, Fertel S, Merli GJ. Use of enoxaparin for the chronically anticoagulated patient before and after procedures. *Am J Cardiol*. 1999;84(4):478-480, A10.
7. Ananthasubramaniam K, Beattie JN, Rosman HS, Jayam V, Borzak S. How safely and for how long can warfarin therapy be withheld in prosthetic heart valve patients hospitalized with a major hemorrhage? *Chest*. 2001;119(2):478-484.
8. Katholi RE, Nolan SP, McGuire LB. The management of anticoagulation during noncardiac operations in patients with prosthetic heart valves: a prospective study. *Am Heart J*. 1978;96(2):163-165.
9. Katholi RE, Nolan SP, McGuire LB. Living with prosthetic heart valves: subsequent noncardiac operations and the risk of thromboembolism or hemorrhage. *Am Heart J*. 1976;92(2):162-167.
10. Ageno W, Garcia D, Libby E, Crowther MA. Managing oral anticoagulant therapy in patients with mechanical heart valves undergoing elective surgery: results of a survey conducted among Italian physicians. *Blood Coagul Fibrinolysis*. 2004;15(8):623-628.
11. Garcia DA, Ageno W, Libby EN, Bibb J, Douketis J, Crowther MA. Perioperative anticoagulation for patients with mechanical heart valves: a survey of current practice. *J Thromb Thrombolysis*. 2004;18(3):199-203.
12. Hylek EM. Trends in anticoagulation management across community-based practices in the United States: the Anticoagulation Consortium to Improve Outcomes Nationally (ACTION) study. *J Thromb Thrombolysis*. 2003;16(1-2):83-86.
13. 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.
14. Kearon C, Hirsh J. Management of anticoagulation before and after elective surgery. *N Engl J Med*. 1997;336(21):1506-1511.
15. Fedi S, Gori AM, Falciani M, et al. Procedure-dependence and tissue factor-independence of hypercoagulability during orthopaedic surgery. *Thromb Haemost*. 1999;81(6):874-878.
16. Kluft C, Verheijen JH, Jie AF, et al. The postoperative fibrinolytic shutdown: a rapidly reverting acute phase pattern for the fast-acting inhibitor of tissue-type plasminogen activator after trauma. *Scand J Clin Lab Invest*. 1985;45(7):605-610.

Anticoagulation

Delivery of Optimized Anticoagulant Therapy: Consensus Statement from the Anticoagulation Forum

David A Garcia, Daniel M Witt, Elaine Hylek, Ann K Wittkowsky, Edith A Nutescu, Alan Jacobson, Stephan Moll, Geno J Merli, Mark Crowther, Laura Earl, Richard C Becker, Lynn Oertel, Amir Jaffer, and Jack E Ansell

OBJECTIVE: To provide recommendations, policies, and procedures pertaining to the provision of optimized anticoagulation therapy designed to achieve desired clinical endpoints while minimizing the risk of anticoagulant-related adverse outcomes (principally bleeding and thrombosis).

STUDY SELECTION AND DATA EXTRACTION: Due to this document's scope, the medical literature was searched using a variety of strategies. When possible, recommendations are supported by available evidence; however, because this paper deals with processes and systems of care, high-quality evidence (eg, controlled trials) is unavailable. In these cases, recommendations represent the consensus opinion of all authors who constitute the Board of Directors of The Anticoagulation Forum, an organization dedicated to optimizing anticoagulation care. The Board is composed of physicians, pharmacists, and nurses with demonstrated expertise and significant collective experience in the management of patients receiving anticoagulation therapy.

DATA SYNTHESIS: Recommendations for delivering optimized anticoagulation therapy were developed collaboratively by the authors and are summarized in 9 key areas: (I) Qualifications of Personnel, (II) Supervision, (III) Care Management and Coordination, (IV) Documentation, (V) Patient Education, (VI) Patient Selection and Assessment, (VII) Laboratory Monitoring, (VIII) Initiation and Stabilization of Warfarin Therapy, and (IX) Maintenance of Therapy. Recommendations are intended to inform the development of care systems containing elements with demonstrated benefit in improvement of anticoagulation therapy outcomes. Recommendations for delivering optimized anticoagulation therapy are intended to apply to all clinicians involved in the care of outpatients receiving anticoagulation therapy, regardless of the structure and setting in which that care is delivered.

CONCLUSIONS: Anticoagulation therapy, although potentially life-saving, has inherent risks. Whether a patient is managed in a solo practice or a specialized anticoagulation management service, a systematic approach to the key elements outlined herein will reduce the likelihood of adverse events. The need for continued research to validate optimal practices for managing anticoagulation therapy is acknowledged.

KEY WORDS: anticoagulant, antithrombotic, thromboembolism, vitamin K antagonist, warfarin.

Ann Pharmacother 2008;42:979-88.

Published Online, 17 Jun 2008, *www.theannals.com*, DOI 10.1345/aph.1L098

Millions of patients receive anticoagulant therapy to prevent or treat thromboembolism. High-quality evidence documenting the benefit of antithrombotic therapy for patients with mechanical heart valves, a history of venous thromboembolism, or atrial fibrillation is abundant.[1-3] However, antithrombotic agents are associated with a risk of bleeding. On death certificates, anticoagulants ranked first in 2003 and 2004 in the number of total mentions of "deaths for drugs causing adverse effects in therapeutic use."[4]

In the US, the Joint Commission has brought significant attention to the safety of antithrombotic agents by challenging hospitals to "reduce the likelihood of patient harm associated with the use of anticoagulation therapy" as 1 of 2 new National Patient Safety Goals for 2008.[5] Warfarin, the only oral anticoagulant available in North America, is notorious for having both a narrow therapeutic index as well as numerous drug and dietary interactions.[6-10] The fear of bleeding complications and the need for frequent blood sampling are among the reasons that oral anticoagulant therapy is underutilized.[11-14]

The safety and effectiveness of both short- and long-term anticoagulation can

Author information provided at the end of the text.

be optimized by a "systematic," evidence-based approach to therapy, often in the context of dedicated anticoagulation management services (AMS).[15,16] The benefits of AMS are well documented.[16,17] However, the majority of anticoagulated patients in North America do not receive care from such services. Thus, recommendations for delivering optimized anticoagulation therapy (OAT) should apply to all clinicians involved in the care of outpatients receiving anticoagulation, regardless of the structure and setting in which that care is delivered.

This document is focused on outpatient care and describes policies and procedures designed to achieve desired clinical endpoints while minimizing the risk of anticoagulant-related adverse outcomes (principally, bleeding and thrombosis). Recommendations in this document are, whenever possible, supported by the best available evidence. However, for some issues, published evidence is inconclusive or unavailable. In all instances, recommendations herein represent the consensus opinion of all authors. We constitute the Board of Directors of the Anticoagulation Forum, an organization dedicated to optimizing anticoagulation care for all patients (www.acforum.org).

Section I: Qualifications of Personnel

1.1 Optimized anticoagulant therapy should be provided by healthcare professionals licensed in a patient-oriented field (eg, medicine, nursing, pharmacy) possessing core competency related to anticoagulation therapy.

COMMENT

Healthcare professionals involved in the management of antithrombotic therapy should be educated in a clinical discipline, trained in patient assessment and care, and licensed in a patient-oriented healthcare field. Technical support personnel (eg, medical assistant, pharmacy technician, nurse technician) may assist in selected aspects of anticoagulation management, including obtaining laboratory test results, scheduling appointments, and other nonclinical duties, but should not be directly involved in patient assessment and therapy management.

Because anticoagulant therapy is complex and associated with substantial risks, additional training is recommended. This training may be obtained in the work environment, through a formal didactic and/or experiential training program, or through self-study.[15] Such additional training, however, should not replace the aforementioned requirements regarding clinical training and licensing necessary to provide patient care. Examples of formal anticoagulant therapy management training programs are listed in Table 1. Core domains of competency for providers of OAT are outlined in Table 2. The National Certification Board of Anticoagulation Providers has been a pioneer in helping US healthcare professionals document (and be recognized for) their expertise in this area (www.ncbap.org/).

Section II: Supervision

2.1 In situations where OAT is provided by a dedicated AMS, a collaborative practice agreement with the healthcare practitioner(s) or organization ultimately responsible for patient care should be established. The collaborative practice agreement should assign day-to-day responsibility for anticoagulation management to AMS personnel and should clearly describe responsibilities, accountability, and job descriptions.

COMMENT

Optimized anticoagulation therapy is provided by a dedicated AMS in numerous healthcare settings, each with unique characteristics and structural elements and influenced by internal and external regulatory requirements as

Table 1. Anticoagulation Therapy Training Programs

Program	Comment
Certified Anticoagulation Provider	provided by the The National Certification Board for Anticoagulation Providers to formally recognize anticoagulation providers meeting educational and patient-care experiential requirements
Research Institute of the American College of Clinical Pharmacy Anticoagulation Training Program	a 4- to 6-wk intensive training program provided through the University of Texas and the Anticoagulation Clinics of North America
American Society of Health-System Pharmacists Foundation Antithrombotic Pharmacotherapy Traineeship	curriculum consists of a self-study program and a 5-day experiential program
University of Southern Indiana College of Nursing and Allied Health Professions Anticoagulant Therapy Management Certificate Program	interactive 6-wk, 40-h Internet certificate program for nurses, pharmacists, and physicians
Lovelace Clinic Foundation Advanced Preceptorship in the Management of Anticoagulation Therapy and Clinical Thrombosis	an in-depth course designed to meet the needs of physicians, nurses, and pharmacists working in or setting up clinics to monitor and coordinate the care of patients on anticoagulation therapy; each conference is limited to 30–40 participants to promote interaction with course faculty and provide networking opportunities

well as state and federal law.[16,18-20] Regardless of the practice setting, the overall AMS supervisory process and administrative matrix should be described with clarity. Furthermore, the roles and responsibilities of each member of the healthcare team involved in providing OAT, including the referring provider, should be clearly defined. Examples of AMS practice guidelines have been published.[21]

Section III: Care Management and Coordination

3.1 Written policies and procedures for the delivery of optimized anticoagulation therapy should be established and approved by the individual who is ultimately responsible for the delivery of anticoagulant care. Policies and procedures should facilitate communication between all parties with a vested interest in the outcomes of anticoagulant therapy.

COMMENT

Policies and procedures serve as a clinical tool and a quality assurance mechanism to reduce variability in the delivery of care.[22] Any individual or dedicated AMS providing OAT should establish policies and procedures that address common and/or controversial issues that may arise (Table 3). Policies and procedures should be reviewed, updated as evidence becomes available, and approved regularly by appropriate committees (eg, a pharmacy and therapeutics or medical executive committee) and should be widely disseminated throughout the organization. These policies and procedures should also include protocols for routine dosing and follow-up determinations and should be available for review within the clinic at all times.

Coordination of anticoagulation therapy requires timely interaction among the anticoagulation providers, referring physicians, surgeons, specialists, dentists, pharmacists, laboratory personnel, skilled nursing facilities, assisted living facilities, and the patients and their caregivers.[23] Communication failures can result in poor patient outcomes.[24] Effective policies and procedures for the delivery of OAT should reduce fragmentation of care by facilitating communication and transitions between healthcare team members with regard to anticoagulation therapy issues. Communication is essential to ensure optimal therapeutic outcomes and should conform to expectations set forth by applicable regulatory agencies (eg, boards of pharmacy, nursing, and medicine). Examples of AMS policies and procedures have been published.[21]

3.2 An efficient system for scheduling and tracking patients should be utilized.

COMMENT

Suboptimal anticoagulant therapy is often attributable to fragmented systems of care.[24] Key components supporting the delivery of OAT can be categorized as: (1) scheduling, (2) testing, (3) decision support, and (4) communication. A tracking system (eg, an electronic database) should be implemented to minimize the possibility that a patient on anticoagulation therapy could be lost to follow-up, even for a brief period.

Section IV: Documentation

4.1 Accurate and easily accessible documentation systems should be used so that information pertinent to anticoagulation therapy can be retrieved in a timely fashion.

COMMENT

Computer software programs specifically designed to manage all aspects of anticoagulation therapy are widely available.[25] It is also possible to adapt existing computer software applications to meet anticoagulation monitoring needs or to use paper forms. The optimal anticoagulation therapy tracking system for a given healthcare environment should be dictated by factors such as the number of patients being monitored and existing information technology resources. For most settings, computerized anticoagu-

Table 2. Core Domains for Competency for Providers of Optimized Anticoagulant Therapy

Applied physiology and pathophysiology of thromboembolic disorders working knowledge regarding the normal physiological processes of hemostasis and thrombosis, and the etiology, risk factors, and clinical manifestations of pathologic thrombus formation
Patient assessment and management knowledge, skills, and competencies to manage and monitor patients on anticoagulant therapy including the ability to assess the efficacy and toxicity of the prescribed anticoagulant treatment, determine whether the therapeutic goals have been achieved, and identify patient-related variables that affect therapy
Patient education ability to provide patient education that is tailored to patients' specific needs to promote safety, enhance adherence, and positively affect clinical outcomes; perform an educational assessment; develop an educational plan; and document the educational activities in the patient's medical record
Applied pharmacology of antithrombotic agents in-depth knowledge regarding the pharmacologic properties of all antithrombotic drugs

Table 3. Anticoagulation Management Issues for Which Established Policies and Procedures May Be Useful

Assessing the risks and benefits of anticoagulation therapy
Documenting patient's understanding of anticoagulation therapy
Indications for anticoagulation therapy
Indication-specific target INR values
Determining the planned duration of anticoagulation therapy
Initiating anticoagulation therapy
Managing therapeutic and nontherapeutic INR values
Determining monitoring intervals for INR and other laboratory parameters pertinent to anticoagulation therapy (eg, complete blood cell counts, urinalysis)
Defining and documenting adverse events (eg, major bleeding, thromboembolism, death)
Defining the mechanism by which missed appointments will be flagged
Establishing a system for the timely reporting of laboratory results
Managing nonadherence to blood tests or clinic visits
Managing transitions between care settings (eg, inpatient to outpatient, inpatient to skilled nursing, outpatient to inpatient)
Defining criteria for discharging patients from a dedicated AMS
Reimbursement procurement
Defining and assessing quality measures
Interrupting anticoagulation for invasive procedures
Managing anticoagulation therapy during pregnancy
Coordination of anticoagulation therapy during travel
Defining eligibility criteria and follow-up requirements for patient self-testing

AMS = anticoagulation management service; INR = international normalized ratio.

lation tracking applications offer increased efficiency.[25] Table 4 contains a list of parameters that may be helpful to include in an anticoagulation tracking database, regardless of the type of documentation system used.

The documentation system should facilitate access to information relevant to quality assessment. It is suggested that parameters such as percent time-in-therapeutic-range, rates of major bleeding, thromboembolic events, and total deaths be recorded. To assess staff and other resource needs, trends in the number of patients managed should also be tracked.

Section V: Patient Education

5.1 The delivery of optimized anticoagulant care should address the educational needs of patients and their caregivers.

COMMENT

Patient safety is enhanced when patients are actively involved in, understand, and take responsibility for their care.[26] Adherence to a

Table 4. Elements of an Anticoagulation Patient-Tracking and Record-Keeping System

Demographic
name
date of birth
sex
contact information for patient and caregivers (eg, phone numbers, home address, email address)
Treatment
indication(s) for anticoagulant therapy
target INR intensity
start date
anticipated/recommended duration of therapy
tablet strength(s) of vitamin K antagonist
risk factors for bleeding and clotting influencing anticoagulation therapy (eg, fall risk, alcoholism, inherited or acquired thrombphilia)
name, dose, route, frequency of administration, and start and stop dates for concomitant medications that could interfere with vitamin K antagonist (prescription and over-the-counter) including herbal products and dietary supplements
chronological flowchart documenting INR results and vitamin K antagonist dosages and other information pertinent to the patient's anticoagulation care
Communication
patient
documentation of patient education processes
copies of all letters sent to patients
documentation of other patient communications (eg, telephone calls, emails, postal letters)
other healthcare practitioners
summaries of all communications with other healthcare practitioners pertaining to anticoagulation therapy
Miscellaneous
complications of anticoagulation therapy (eg, bleeding, thromboembolism)
other pertinent laboratory values (eg, hemoglobin, hematocrit, urinalysis, fecal occult blood screening)
missed appointments
use of anticoagulants other than vitamin K antagonist (eg, unfractionated heparin, low-molecular-weight heparin, fondaparinux)
plans for interrupting anticoagulation therapy for invasive procedures

INR = international normalized ratio.

plan of care and the stability of anticoagulant effect, as measured by the international normalized ratio (INR), are improved when this is achieved.[27,28] Knowledge of anticoagulation therapy can be effectively imparted through face-to-face interactions and the use of written materials and other audiovisual resources to review and reinforce the educational process.[29] An approach to the learning process based on established models of education may be more likely to improve a patient's knowledge level compared with ad hoc programs.[30]

A knowledge assessment tool may help the clinician to assess an individual patient's educational needs.[31] Written materials at an appropriate reading level should be provided and, when possible, in the patient's native language. Local health literacy rates (a significant concern in many parts of the US) should be considered when patient educational materials are developed.[32] Important aspects of patient education related to anticoagulation therapy are summarized in Table 5.

Section VI: Patient Selection and Assessment

6.1 Optimized anticoagulant therapy should be instituted only after careful consideration of the risk and benefit for an individual patient.

COMMENT

The ability to deliver OAT is highly dependent on patient selection, vigilant INR monitoring, and evidence-based treatment recommendations.[1,33-36] The initial patient assessment should include a comprehensive medical history; family history of bleeding and/or clotting disorders; medications (including dietary supplements and over-the-counter drugs); social, lifestyle, and employment profile; health beliefs and attitudes; level of understanding; health literacy; personal health motivation; and healthcare resources. Risk factors for vitamin K antagonist–associated bleeding have been published.[37,38] Patients and/or their caregivers should be involved in the discussion of the risks and benefits associated with anticoagulation therapy and should agree with the decision as to whether to initiate/ continue therapy.

6.2 The appropriateness of a treatment plan for any individual patient should be periodically reviewed throughout the course of therapy.

COMMENT

A thorough assessment of the various factors that influence warfarin dosing requirements (eg, diet, disease, other medications, alcohol use, adherence) should be completed at all routine patient visits. Since a patient's risk of thrombosis and bleeding can change over time, the indication, intensity, and length of anticoagulation therapy should be reevaluated periodically.[39] Ongoing reassessment will also allow the treating clinician(s) to apply new therapies, algorithms, or guidelines that may be developed.

Section VII: Laboratory Monitoring

7.1 Optimized anticoagulation therapy should incorporate regular laboratory monitoring of anticoagulant effect. Vita-

Table 5. Important Aspects of Anticoagulation Therapy Patient Education

Indicate the reason for initiating anticoagulation therapy and how it relates to clot formation.
Review the name of anticoagulant drug(s) (generic and trade) and discuss how they work to reduce the risk of clotting complications.
Discuss the potential duration of therapy.
Explain the meaning and significance of the INR.
Explain the need for frequent INR testing and target INR values appropriate for the patient's treatment.
Discuss the narrow therapeutic index of warfarin and emphasize the importance of regular monitoring as a way to minimize the risk of bleeding/ thrombosis.
Describe the common signs/symptoms of bleeding and what to do if they occur.
Describe the common signs/symptoms of clotting complications and what to do if they occur.
Outline precautionary measures to minimize the risk of trauma or bleeding.
Discuss the influence of dietary vitamin K use on the effects of vitamin K antagonists.
Discuss potential drug interactions (prescription, over-the-counter, herbal) and what to do when normal medication regimens change.
Discuss the need to avoid or limit alcohol consumption.
Explain the need for birth control measures for women of child-bearing age.
Review the importance of notifying all healthcare providers (eg, physicians, dentists) of the use of anticoagulation therapy.
Review the importance of notifying the anticoagulation provider when dental, surgical, or invasive procedures and hospitalization are scheduled.
Explain when to take anticoagulant medications and what to do if a dose is missed.
Discuss the importance of carrying identification (ID card, medical alert bracelet/necklace).
Document the fact that education of the patient (or caregiver) has occurred.

ID = identification; INR = international normalized ratio.

min K antagonists should be monitored with use of the prothrombin time test and reported as an INR.

COMMENT

Unique preanalytic, analytic, and postanalytic sources of error may, as with all laboratory tests, affect prothrombin time results.[40,41] Even when all of these variables are tightly controlled, there remains a clinically significant amount of variability between different test systems, depending on the specific coagulometer and thromboplastin combination utilized.[42-44] The reproducibility of results when repeated testing is performed on the same test system is quite precise, with a coefficient of variation generally below 5%. Replicate testing of the same sample on multiple different test systems results in a much greater degree of variation, and this variation increases significantly with higher intensity of anticoagulation.[45,46] Despite this variation between different test systems, the prothrombin time (and its derivative, the INR) has been shown to correlate with important outcomes in multiple clinical trials.[6-8,47,48]

The INR is a standardization method that attempts to minimize differences between thromboplastin reagents through a calibration process in which all commercial thromboplastins are compared with an International Reference Preparation (IRP) maintained by the World Health Organization (WHO).[49] The INR method is not perfect in correcting for differences among different laboratories utilizing different thromboplastin reagents, but it does reduce the variation among different laboratories and provides clinically useful results.[50,51]

7.2 Prothrombin time testing for optimized anticoagulation therapy should be performed on either plasma samples in a clinical laboratory or on whole blood capillary (fingerstick) samples utilizing point-of-care devices.

COMMENT

Both approaches have been validated and both provide results equivalent to results obtained with WHO IRP preparations. Both plasma (venipuncture) and whole blood (fingerstick) methods of prothrombin time testing have been used for decision-making in anticoagulant-related clinical trials.[52-54]

7.3 Prothrombin time testing for optimized anticoagulation therapy should be performed by professional laboratory staff, professional clinical staff, or properly trained patients or caregivers.

COMMENT

Laboratory testing has traditionally been performed in a clinical laboratory by trained laboratory professionals. The development of whole blood prothrombin time testing has more recently allowed for the testing to move outside of the clinical laboratory. Multiple studies have validated that not only nonlaboratory medical professionals, but also properly trained patients, are capable of performing reliable prothrombin time testing. The bulk of the data suggest that, for properly selected patients, self-testing (at home) is cost-effective and leads to outcomes at least as good as those achieved with standard INR testing (at a clinical laboratory or in a clinic).[52,55] Barriers to widespread adoption of patient self-testing in the US include: (1) the lack of a single, evidence-based approach to identifying eligible patients, (2) reluctance on the part of many third-party payers to fund the machines and the test strips, and (3) the absence (in many primary care settings) of a well-developed system with which self-testing patients can be identified, educated, and have their follow-up ensured. Whether it is performed by the patient (at home) or by a healthcare professional (in a medical office), point-of-care testing offers efficiency for the clinician and eliminates any potential for delay between INR measurement and patient notification of results.

Section VIII: Initiation and Stabilization of Warfarin Therapy

8.1 The initiation of optimized anticoagulation therapy should use a systematic, evidence-based approach.

COMMENT

The initiation of OAT should ensure that therapeutic concentrations of anticoagulant medications are achieved in a timely manner and that the risk of supratherapeutic INR values is minimized. Various approaches to achieving this goal are outlined in evidence-based guidelines and the medical literature and should be used in the development of systems for the initiation of OAT.[56-60] Clinicians should consider patient-specific factors such as age, weight, height, concomitant medications, and comorbidities when deciding on the starting doses of anticoagulant medications. Irrespective of the starting dose used, INR values should be monitored at least 2–3 times per week for the first 7–10 days (or until a stable dose is achieved) of vitamin K antagonist therapy.[56]

Although the presence of certain polymorphisms in the genes for CYP2C9 and vitamin K epoxide reductase complex subunit 1 is associated with lower maintenance doses, the role of pharmacogenetic testing in clinical practice remains uncertain. Several clinical trials designed to test the hypothesis that pharmacogenetic testing will improve patient care are ongoing. At this time, however, we do not believe that there is sufficient evidence of benefit to recommend routinely genotyping patients who initiate vitamin K antagonist treatment.[61-63]

Patients being started on vitamin K antagonist treatment often require concomitant unfractionated heparin, low-molecular-weight heparin (LMWH), or synthetic pentasaccharide (fondaparinux) during vitamin K antagonist initiation.[56] Healthcare professionals supervising initiation of vitamin K antagonist treatment should define the answers to questions such as: What laboratory parameters should be checked and how often? When should "overlap" heparin/LMWH/fondaparinux therapy be discontinued?

Section IX: Maintenance of Therapy

9.1 The delivery of OAT should use a systematic process for longitudinal patient assessment, adjustment of anticoagulant drug doses, and scheduling of follow-up laboratory monitoring.

COMMENT

Follow-up evaluation during OAT should document changes in medication, health status, diet, and adherence. Patients should also be assessed regularly for signs and symptoms of bleeding or clotting complications. Standardization of follow-up procedures using checklists or flow diagrams may increase the consistency of care.[64] For patients on a stable dose of a vitamin K antagonist, individual circumstances, such as medication changes, concurrent illness, or unexplained INR instability, will dictate the interval between follow-up assessments. However, current guidelines indicate that even patients with repeatedly therapeutic levels of anticoagulation should undergo INR measurement every 4 weeks.[56]

Validated algorithms for adjusting vitamin K antagonist doses should be incorporated into operating procedures. Evidence-based guidelines should be used to establish a systematic approach to responding to extreme INR values (eg, >4.5 and <1.5).[56] Likewise, a systematic approach that incorporates pharmacokinetic and pharmacodynamic principles should be employed to determine the interval between INR tests that maximizes the amount of time that anticoagulant concentrations are maintained within their therapeutic range.

9.2 The delivery of optimized anticoagulation therapy should utilize a systematic approach to the elective interruption and resumption of anticoagulant therapy for elective invasive procedures.

COMMENT

Patients receiving a vitamin K antagonist may require temporary interruption of anticoagulant therapy to minimize bleeding risk associated with invasive procedures. The risk of excessive or uncontrolled bleeding associated with the procedure should be carefully weighed against the potential for recurrent thromboembolism associated with the interruption in anticoagulation therapy.[65-67] Although no high-quality evidence to guide perioperative anticoagulation decisions exists, local (or institutional) standards regarding protocols, communication with interventionists, and patient education will reduce inconsistency when patients require invasive procedures. Both the person responsible for managing anticoagulation therapy and the person performing the invasive procedure should be in agreement regarding the anticoagulation therapy plan. Consensus guidelines, although based on evidence of limited quality, addressing this common clinical situation have been published (ref *Chest*, ACC/AHA, *International Angiology*).[34-36,56]

9.3 The delivery of optimized anticoagulation therapy should use a systematic approach in management and documentation of unexpected events (eg, bleeding, clotting, other potential anticoagulation-related adverse effects, or medical problems not related to anticoagulant therapy).

COMMENT

Patients experiencing unexpected adverse events should be triaged and managed in a setting where the required care can be provided in a timely manner. Preferred interventions for the prompt reduction of the INR in bleeding patients (eg, infusions of fresh-frozen plasma, prothrombin complex concentrates, or recombinant factor VIIa, along with vitamin K) should be developed collaboratively with emergency care providers and based on available evidence.[56] The severity and location of the bleeding and the level of the INR should influence the approach and choice of a reversal agent. Policies should also be in place for managing patients with subtherapeutic INR results and/or thromboembolic events in a timely manner. As with patients who experience (or are at risk for) bleeding events, the plan for those presenting with a low INR or signs/symptoms of a thrombotic event will be dictated by clinical circumstances such as the underlying risk of thrombosis and the length of time during which the INR has been subtherapeutic. Systems should be developed to facilitate continuity of care when patients first seek medical attention in an emergency department. Any treatment rendered should be documented and communicated in a timely fashion to the person managing anticoagulation therapy.

Summary

Anticoagulation therapy, although potentially life-saving, has inherent risks. Whether a patient is managed in a solo practice or a specialized AMS, a systematic approach to key elements will reduce the likelihood of adverse events. The guidelines in this article are intended to help healthcare providers at the point of delivery to optimize anticoagulation therapy. Even as new anticoagulant medications emerge, the

principles of patient selection, provider education and training, interruption of treatment for invasive procedures, and careful follow-up are likely to remain relevant.

David A Garcia MD, Associate Professor, University of New Mexico School of Medicine; Medical Director, Anticoagulation Clinic, University of New Mexico Health Sciences Center, University of New Mexico, Albuquerque, NM

Daniel M Witt PharmD, Clinical Assistant Professor, School of Pharmacy, University of Colorado; Manager, Clinical Pharmacy Services, Kaiser Permanente Colorado, Denver, CO

Elaine Hylek MD, Associate Professor, School of Medicine, Boston University; Director, Anticoagulation Clinic, Boston Medical Center, Boston, MA

Ann K Wittkowsky PharmD CACP FASHP FCCP, Clinical Professor, School of Pharmacy, University of Washington; Director, Anticoagulation Services, University of Washington Medical Center, Seattle, WA

Edith A Nutescu PharmD FCCP, Clinical Associate Professor of Pharmacy Practice; Director, Antithrombosis Center, College of Pharmacy & Medical Center, The University of Illinois at Chicago, Chicago, IL

Alan Jacobson MD, Director, Anticoagulation Services; Associate Chief of Staff for Research, Loma Linda Veterans Affairs Medical Center, Loma Linda, CA

Stephan Moll MD, Associate Professor, School of Medicine, University of North Carolina, Chapel Hill, NC

Geno J Merli MD FACP, Professor of Medicine, Jefferson Medical College, Thomas Jefferson University; Senior Vice President and Chief Medical Officer, Thomas Jefferson University Hospital, Philadelphia, PA

Mark Crowther MD MSc FRCPC, Professor and Academic Division Director, Hematology and Thromboembolism, McMaster University; Director of Laboratory Hematology, Hamilton Regional Laboratory Medicine Program; Head of Service, Hematology, St. Joseph's Healthcare, Hamilton, Ontario, Canada

Laura Earl BSN CACP, Practicing Nurse, University of New Mexico, Albuquerque

Richard C Becker MD, Professor of Medicine, School of Medicine, Duke University; Director, Duke Cardiovascular Thrombosis Center, Duke Clinical Research Institute, Durham, NC

Lynn Oertel MS ANP CACP, Clinical Nurse Specialist, Anticoagulation Management Services, Massachusetts General Hospital, Boston, MA

Amir Jaffer MD, Division Chief of Hospital Medicine, University of Miami, Miami, FL

Jack E Ansell MD, Chair, Department of Medicine, Lenox Hill Hospital, New York, NY

Reprints: Dr. Garcia, MSC10 5550, University of New Mexico Health Sciences Center, Albuquerque, NM 87131, fax 505/272-0358, DAVGarcia@salud.unm.edu

Dr. Garcia has served as a consultant to ARYx Therapeutics, AstraZeneca, Boehringer Ingelheim, Bristol-Myers Squibb, CSL Behring, Roche Diagnostics, and Sanofi-Aventis and has received research funding from Bristol-Myers Squibb and CSL Behring.

Dr. Hylek has served as a consultant to Bristol-Myers Squibb and has received research funding from AstraZeneca and Bristol-Myers Squibb.

Dr. Wittkowsky has served as a consultant for Boehringer Ingelheim and has served on the speaker's bureau for Sanofi-Aventis.

Dr. Jacobson has served as a consultant for Boehringer Ingelheim, Bristol-Myers Squibb, Farallon Medical, GlaxoSmithKline, Hemosense, Inverness Medical, ITC, Roche Diagnostics, and Sanofi-Aventis.

Dr. Moll has served as a consultant for Bayer, Bristol-Myers Squibb, Inverness, ITC, and Roche Diagnostics.

Dr. Merli has served as a consultant for Bayer, Bristol-Myers Squibb, and Sanofi-Aventis and has received research funding from Bayer, Bristol-Myers Squibb, and Sanofi-Aventis.

Ms. Earl has served as a consultant for ARYx Therapeutics, and Roche Diagnostics and has served on the speaker's bureau for Roche Diagnostics.

Dr. Becker has received research funding from AstraZeneca, Bayer, Bristol-Myers Squibb, Johnson & Johnson, Medicines Company, Regado Biosciences, and Schering-Plough Research Institute.

Ms. Oertel has served on the speaker's bureau for GlaxoSmithKline.

Dr. Jaffer has served as a consultant for AstraZeneca, Boehringer Ingelheim, Roche Diagnostics, and Sanofi-Aventis, served on the speaker's bureau for Sanofi-Aventis, and received research funding from AstraZeneca.

Dr. Ansell has served as a consultant for ARYx Therapeutics, Bayer, Bristol-Myers Squibb, Hemosense, Inverness, ITC, and Roche Diagnostics and has served on the speaker's bureau for Bayer, Bristol-Myers Squibb, Boehringer Ingelheim, Roche Diagnostics, and Sanofi-Aventis.

References

1. Buller HR, Agnelli G, Hull RD, Hyers TM, Prins MH, Raskob GE. Antithrombotic therapy for venous thromboembolic disease: the Seventh ACCP Conference on Antithrombotic and Thrombolytic Therapy. Chest 2004;126(suppl):401S-28S.
2. Connolly S, Pogue J, Hart R, 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 2006;367:1903-12.
3. Stein PD, Alpert JS, Bussey HI, Dalen JE, Turpie AG. Antithrombotic therapy in patients with mechanical and biological prosthetic heart valves. Chest 2001;119(suppl):220S-7S.
4. Wysowski DK, Nourjah P, Swartz L. Bleeding complications with warfarin use: a prevalent adverse effect resulting in regulatory action. Arch Intern Med 2007;167:1414-9.
5. Joint Commission. www.jointcommission.org/PatientSafety/NationalPatientSafetyGoals/08_hap_npsgs.htm (accessed 2008 May 28).
6. 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;349:1019-26.
7. Hylek EM, Singer DE. Risk factors for intracranial hemorrhage in outpatients taking warfarin. Ann Intern Med 1994;120:897-902.
8. 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;335:540-6.
9. Juurlink DN. Drug interactions with warfarin: what clinicians need to know. CMAJ 2007;177:369-71.
10. Khan T, Wynne H, Wood P, et al. Dietary vitamin K influences intra-individual variability in anticoagulant response to warfarin. Br J Haematol 2004;124:348-54.
11. Bungard TJ, Ghali WA, Teo KK, McAlister FA, Tsuyuki RT. Why do patients with atrial fibrillation not receive warfarin? Arch Intern Med 2000;160:41-6.
12. Cohen N, Almoznino-Sarafian D, Alon I, et al. Warfarin for stroke prevention still underused in atrial fibrillation: patterns of omission. Stroke 2000;31:1217-22.
13. Fang MC, Stafford RS, Ruskin JN, Singer DE. National trends in antiarrhythmic and antithrombotic medication use in atrial fibrillation. Arch Intern Med 2004;164:55-60.
14. Partington SL, Abid S, Teo K, Oczkowski W, O'Donnell MJ. Pre-admission warfarin use in patients with acute ischemic stroke and atrial fibrillation: the appropriate use and barriers to oral anticoagulant therapy. Thromb Res 2007;120:663-9.
15. Dager WE, Branch JM, King JH, et al. Optimization of inpatient warfarin therapy: impact of daily consultation by a pharmacist-managed anticoagulation service. Ann Pharmacother 2000;34:567-72.
DOI 10.1345/aph.18192

16. 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;127:1515-22.
17. Locke C, Ravnan SL, Patel R, Uchizono JA. Reduction in warfarin adverse events requiring patient hospitalization after implementation of a pharmacist-managed anticoagulation service. Pharmacotherapy 2005;25: 685-9.
18. 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;158:1641-7.
19. Ernst ME, Brandt KB. Evaluation of 4 years of clinical pharmacist anticoagulation case management in a rural, private physician office. J Am Pharm Assoc 2003;43:630-6.
20. Holden J, Holden K. Comparative effectiveness of general practitioner versus pharmacist dosing of patients requiring anticoagulation in the community. J Clin Pharm Ther 2000;25:49-54.
21. Oertel L. Personnel needs and division of labor. In: Ansell JE, Oertel L, Wittkowsky AK, eds. Managing oral anticoagulation therapy. New York: Aspen, 2003:4:1-4:30.
22. Tillman DJ, Charland SL, Witt DM. Effectiveness and economic impact associated with a program for outpatient management of acute deep vein thrombosis in a group model health maintenance organization. Arch Intern Med 2000;160:2926-32.
23. McCormick W. Medical–legal implications of anticoagulation therapy. In: Ansell J, Oertel L, Wittkowsky A, eds. Managing oral anticoagulation therapy. 2nd ed. St. Louis: Wolters Kluwer Health, Inc., 2005:14:1-14:7.
24. Ryan E. Risk management and anticoagulation therapy. In: Ansell J, Oertel L, Wittkowsky A, eds. Managing oral anticoagulation therapy. 2nd ed. St. Louis: Wolters Kluwer Health, Inc, 2005:9A-1, 1-11.
25. Oertel L, Mungall D. Software applications in anticoagulation management. In: Ansell J, Oertel L, Wittkowsky A, eds. Managing oral anticoagulation therapy. 2nd ed. St. Louis: Wolters Kluwer Health, Inc, 2005:5A-2, 1-11.
26. Beyth RJ, Quinn L, Landefeld CS. A multicomponent intervention to prevent major bleeding complications in older patients receiving warfarin. A randomized, controlled trial. Ann Intern Med 2000;133:687-95.
27. Saligari E, Belle L, Berry C, et al. [Evaluation of an education program of patients undergoing oral anticoagulation treatment] French. Ann Cardiologie Angeiol (Paris) 2003;52:297-301.
28. Barcellona D, Contu P, Marongiu F. Patient education and oral anticoagulant therapy. Haematologica 2002;87:1081-6.
29. Stone S, Holden A, Knapic N, Ansell J. Comparison between videotape and personalized patient education for anticoagulant therapy. J Fam Pract 1989;29:55-7.
30. Newall F, Monagle P, Johnston L. Patient understanding of warfarin therapy: a review of education strategies. Hematology (Amsterdam) 2005; 10:437-42.
31. Zeolla MM, Brodeur MR, Dominelli A, Haines ST, Allie ND. Development and validation of an instrument to determine patient knowledge: the oral anticoagulation knowledge test. Ann Pharmacother 2006;40:633-8. Epub 21 Mar 2006. DOI 10.1345/aph.1G562
32. Estrada CA, Martin-Hryniewicz M, Peek BT, Collins C, Byrd JC. Literacy and numeracy skills and anticoagulation control. Am J Med Sci 2004; 328:88-93.
33. Albers GW, Amarenco P, Easton JD, Sacco RL, Teal P. Antithrombotic and thrombolytic therapy for ischemic stroke: the Seventh ACCP Conference on Antithrombotic and Thrombolytic Therapy. Chest 2004; 126(suppl):483S-512S.
34. Salem DN, Stein PD, Al-Ahmad A, et al. Antithrombotic therapy in valvular heart disease—native and prosthetic: the Seventh ACCP Conference on Antithrombotic and Thrombolytic Therapy. Chest 2004;126(suppl):457S-82S.
35. Bonow RO, Carabello BA, Kanu C, et al. ACC/AHA 2006 guidelines for the management of patients with valvular heart disease: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (writing committee to revise the 1998 Guidelines for the Management of Patients with Valvular Heart Disease): developed in collaboration with the Society of Cardiovascular Anesthesiologists: endorsed by the Society for Cardiovascular Angiography and Interventions and the Society of Thoracic Surgeons. Circulation 2006; 114:e84-231.
36. 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:e257-354.
37. 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;105:91-9.
38. 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;130:1390-6.
39. Kearon C. Long-term management of patients after venous thromboembolism. Circulation 2004;110(9 suppl 1):I10-8.
40. 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;131:816-22.
41. Johnson SG, Witt DM, Eddy TR, Delate T. Warfarin and antiplatelet combination use among commercially insured patients enrolled in an anticoagulation management service. Chest 2007;131:1500-7.
42. van Rijn JL, Schmidt NA, Rutten WP. Correction of instrument- and reagent-based differences in determination of the international normalized ratio (INR) for monitoring anticoagulant therapy. Clin Chem 1989; 35:840-3.
43. 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;10:215-27.
44. Woodhams BJ, Klein N, Harz D, Rose M, Ruiz JA. A study of the variability seen in the international normalized ratio obtained using different sensitivity thromboplastin reagents on different instrument types. Blood Coagul Fibrinolysis 1999;10:423-7.
45. Horsti J, Uppa H, Vilpo JA. Poor agreement among prothrombin time international normalized ratio methods: comparison of seven commercial reagents. Clin Chem 2005;51:553-60.
46. Olson JD, Brandt JT, Chandler WL, et al. Laboratory reporting of the international normalized ratio: progress and problems. Arch Pathol Lab Med 2007;131:1641-7.
47. Kearon C, Ginsberg JS, Kovacs MJ, et al. 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.
48. Ridker PM, Goldhaber SZ, Danielson E, et al. Long-term, low-intensity warfarin therapy for the prevention of recurrent venous thromboembolism (comment). N Engl J Med 2003;348:1425-34.
49. Loeliger EA, van den Besselaar AM, Lewis SM. Reliability and clinical impact of the normalization of the prothrombin times in oral anticoagulant control. Thromb Haemos 1985;53:148-54.
50. Hirsh J, Poller L. The international normalized ratio. A guide to understanding and correcting its problems. Arch Intern Med 1994;154:282-8.
51. Pi DW, Raboud JM, Filby C, Carter CJ. Effect of thromboplastin and coagulometer interaction on the precision of the international normalised ratio. J Clin Pathol 1995;48:13-7.
52. 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;367:404-11.
53. 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;155:1861-7.
54. Menendez-Jandula B, Souto JC, Oliver A, et al. Comparing self-management of oral anticoagulant therapy with clinic management: a randomized trial. Ann Intern Med 2005;142:1-10.

55. Regier DA, Sunderji R, Lynd LD, Gin K, Marra CA. Cost-effectiveness of self-managed versus physician-managed oral anticoagulation therapy. CMAJ 2006;174:1847-52.
56. Ansell J, Hirsh J, Poller L, Bussey H, Jacobson A, Hylek E. The pharmacology and management of the vitamin K antagonists: the Seventh ACCP Conference on Antithrombotic and Thrombolytic Therapy. Chest 2004;126(suppl):204S-33S.
57. 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;138:714-9.
58. Crowther MA, Ginsberg J, Kearon C, et al. A randomized trial comparing 5 mg and 10 mg warfarin loading doses. Arch Intern Med 1999; 159:46-8.
59. O'Connell MB, Kowal PR, Allivato CJ, Repka TL. Evaluation of warfarin initiation regimens in elderly inpatients. Pharmacotherapy 2000; 20:923-30.
60. 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; 118:137-42.
61. Joffe HV, Xu R, Johnson FB, Longtine J, Kucher N, Goldhaber SZ. Warfarin dosing and cytochrome P450 2C9 polymorphisms. Thromb Haemos 2004;91:1123-8.
62. Gage BF, Eby C, Milligan PE, Banet GA, Duncan JR, McLeod HL. Use of pharmacogenetics and clinical factors to predict the maintenance dose of warfarin. Thromb Haemos 2004;91:87-94.
63. Anderson JL, Horne BD, Stevens SM, et al. Randomized trial of genotype-guided versus standard warfarin dosing in patients initiating oral anticoagulation. Circulation 2007;116:2563-70.
64. Oertel L. Managing maintenance therapy. In: Ansell J, Oertel L, Wittkowsky A, eds. Managing oral anticoagulation therapy. 2nd ed. St. Louis: Wolters Kluwer Health, Inc., 2005, 4B-3, 1-5.
65. Douketis JD, Johnson JA, Turpie AG. Low-molecular-weight heparin as bridging anticoagulation during interruption of warfarin: assessment of a standardized periprocedural anticoagulation regimen. Arch Intern Med 2004;164:1319-26.
66. Dunn AS, Turpie AG. Perioperative management of patients receiving oral anticoagulants: a systematic review. Arch Intern Med 2003;163:901-8.
67. Kovacs MJ, Kearon C, Rodger M, et al. Single-arm study of bridging therapy with low-molecular-weight heparin for patients at risk of arterial embolism who require temporary interruption of warfarin. Circulation 2004;110:1658-63.

Fournir une Thérapie Anticoagulante Optimale: Prise de Position du Forum d'Anticoagulation.

DA Garcia, DM Witt, E Hylek, AK Wittkowsky, EA Nutescu, A Jacobson, S Moll, GJ Merli, M Crowther, L Earl, RC Becker, L Oertel, A Jaffer, et JE Ansell

Ann Pharmacother 2008;42:979-88.

RÉSUMÉ

OBJECTIF: Donner des recommandations, des politiques, et des procédures concernant l'offre d'un service d'anticoagulation optimisé visant l'atteinte de résultats cliniques tout en minimisant le risque de résultats indésirables liés à la thérapie anticoagulante (saignements et thromboses).

SÉLECTION DES ÉTUDES ET DE L'INFORMATION: Considérant l'envergure du document, la littérature médicale a été scrutée à l'aide de différentes stratégies. Lorsque possible, les recommandations furent supportées par les évidences disponibles. Cependant, parce que ce manuscrit fait référence aux processus et systèmes de soins, des évidences de haute qualité (telle une étude randomisée) ne sont pas disponibles. Dans ces cas, les recommandations représentent un consensus d'opinions des auteurs participant au Conseil des Directeurs du Forum d'Anticoagulation, une organisation dédiée à optimiser les soins en anticoagulation. Ce conseil est composé de médecins, pharmaciens, et infirmières ayant démontré une expertise et une expérience significative dans le traitement de ces patients.

RÉSUMÉ: Les recommandations pour fournir une thérapie anticoagulante optimisée furent développées en collaboration avec les auteurs et se résument en 9 éléments clés: (1) Qualification du personnel, (2) Supervision, (3) Gestion des soins et sa coordination, (4) Documentation, (5) Education du patient, (6) Sélection du patient et évaluation, (7) Monitorage de laboratoire, (8) Initiation et stabilisation de la thérapie, et (9) Maintien de la thérapie. Les recommandations veulent favoriser le développement de systèmes de soins dont les éléments ont démontré des bénéfices dans l'amélioration des résultats des patients anticoagulés. Les recommandations pour dispenser une thérapie optimisée de soins de santé s'appliquent à tous les cliniciens impliqués dans les soins à ces patients, peu importe la structure et l'endroit où les soins sont donnés.

CONCLUSIONS: La thérapie anticoagulante, bien que salutaire, comporte des risques. Lorsqu'un patient est suivi par une personne seule ou par un service spécialisé en anticoagulation, une approche systématique comportant les 9 éléments cités permettra de réduire les risques inhérents à cette thérapie. La recherche continue pour dispenser une thérapie anticoagulante optimale est nécessaire et le besoin reconnu.

Traduit par Marc M Perreault

Update in Addiction Medicine for the Primary Care Clinician

Adam J. Gordon[1,2], David A. Fiellin[3], Peter D. Friedmann[4], Marc N. Gourevitch[5], Kevin L. Kraemer[2], Julia H. Arnsten[6], Richard Saitz[7,8] and for the Society of General Internal Medicine's Substance Abuse Interest Group

[1]Center for Health Equity Research and Promotion, VA Pittsburgh Healthcare System and Mental Illness Research Education and Clinical Center, VA Pittsburgh Healthcare System, Pittsburgh, PA, USA; [2]Center for Research on Health Care, University of Pittsburgh, Pittsburgh, PA, USA; [3]Yale University School of Medicine, New Haven, CT, USA; [4]Center on Systems, Outcomes and Quality in Chronic Disease and Rehabilitation, Providence VA Medical Center, Division of General Internal Medicine, Rhode Island Hospital, Alpert Medical School of Brown University, Providence, RI, USA; [5]New York University School of Medicine, New York, NY, USA; [6]Albert Einstein College of Medicine and Montefiore Medical Center, New York, NY, USA; [7]Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Boston Medical Center and Boston University School of Medicine, Boston, MA, USA; [8]Youth Alcohol Prevention Center and Department of Epidemiology, Boston University School of Public Health, Boston, MA, USA.

KEY WORDS: substance-related disorders; primary health care; review literature.
J Gen Intern Med 23(12):2112–6
DOI: 10.1007/s11606-008-0806-7

INTRODUCTION

The United States Preventive Services Task Force recommends that primary care clinicians assume a major role in screening, identification, treatment, and referral to treatment of unhealthy alcohol and other drug (AOD) use—the spectrum from use that risks health consequences to AOD disorders (abuse and dependence)—in generalist settings.[1] In the United States, nicotine dependence, alcohol use, and drug use are the first, third, and ninth leading causes, respectively, of preventable deaths.[2] Despite the harmful effects of addiction and improved options for office-based treatments and referral, not all primary care clinicians routinely address AOD use in their patients.

The objectives of this paper are to identify and examine important recent advances in addiction medicine that have implications for primary care clinicians and that emphasize primary care clinicians' role in the identification, treatment and/or referral of patients with addictions. We conducted an electronic database (PubMed) search to systematically identify recent (June 1, 2006 to January 1, 2008) human subject, English language, peer-reviewed, research publications that are relevant to generalist care for patients with addiction disorders. We also surveyed the publications that were reviewed by a NIH-funded newsletter that, in an attempt to identify articles that address the health impact of alcohol and drugs, systematically reviews the core general medical, infectious disease, public health, and addiction subspecialty journals.[3] Similar to our prior review,[4] authors (A.G., D.F., R.S.) were provided a title listing of articles with addiction-related key words within the reference time frame, and then secondary searches and consensus deliberations were used to identify articles that may impact the care provided by primary care clinicians in the categories of 1) alcohol use and disorders and 2) opioid use and dependence. Articles were categorized as impacting primary care clinicians if they studied primary care settings or could impact such settings and had practice-changing findings or implications.

Received July 30, 2008
Revised August 24, 2008
Accepted August 28, 2008
Published online October 2, 2008

ALCOHOL USE AND DISORDERS

Simplifying Alcohol Assessment: Two Questions to Identify Alcohol Use Disorders[5]

In this study,[5] the investigators developed a simple assessment for alcohol use disorders with data from the cases (1,522 injured patients seen in an emergency department) of a case-control study, the Missouri Injury Study (MIS). They validated the assessment with data from three cross-sectional samples: 1) the controls (1,124 non-injured adults responding to a telephone survey) from the MIS, 2) a primary care sample (n=623) from the Vital Signs Screening Project (VSSP), and 3) a nationally representative sample of U.S. adults (n=26,946) from the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC).

The investigators identified two criteria that were predictive of alcohol use disorders: 1) recurrent drinking in physically hazardous situations and 2) drinking more or for longer than intended. In the developmental sample, presence of either of the two criteria had a sensitivity of 96% and a specificity of 85% for a current alcohol use disorder. Among all subjects in the three validation samples, the presence of either of the two criteria had a sensitivity of 72% to 94% and a specificity of 80% to 95% for a current alcohol use disorder. Among screen-positive (>4 drinks in one day for women or >5 drinks in one day for men at least once over the past three months for MIS and VSSP; ≥5 or more drinks in a single day over the past 12 months for NESARC) subjects in the three validation samples, the presence of either of the two criteria had a sensitivity of 77% to 95% and a specificity of 62% to 86% for a current alcohol use disorder.

Although the two criteria—recurrent drinking in physically hazardous situations and drinking more or for longer than intended—demonstrated good sensitivity and specificity for alcohol use disorders in this retrospective secondary data analysis study, it is unclear how the relevant questions should be worded for use in primary care settings and how they would perform outside the context of an extensive research assessment. The investigators recommend the items be tested prospectively in practice settings. After development into an easy-to-use format with prospective validation, this approach could be a brief and efficient way to assess alcohol use disorders in primary care patients with an initial positive screen. Being able to quickly sort out the severity of unhealthy alcohol use in these settings could decrease at least one barrier to widespread dissemination of alcohol screening.

Brief Interventions Effective in Primary Care[6]

In numerous randomized trials, brief counseling interventions, including feedback, advice, and goal-setting done in an empathic manner, have efficacy for nondependent unhealthy alcohol use. But as with many efficacy trials, these studies have had numerous exclusions and the interventions have often been very carefully implemented or extensive, raising questions about reproducibility. Thus, effectiveness of alcohol brief intervention in real-world primary care settings is less clear. To assess the effectiveness of brief interventions, investigators performed a systematic review and meta-analysis of randomized trials done with patients presenting to primary care for reasons other than for alcohol treatment.[6] Interventions in these studies lasted from 7.5 to 60 minutes over 1 to 5 contacts. The researchers scored the trials on the presence or absence of features of efficacy and effectiveness trials. They identified 22 trials with 7619 subjects.

Brief intervention subjects drank (mean difference) 38 grams (about 3–4 standard drinks[7]) less per week than did control subjects at one year (95% CI, −54 to −23 grams). There was no significant difference in reductions between "effectiveness" and "efficacy" trials. In studies characterized as effectiveness trials, the difference in consumption was 28 g (95% CI, −48 to −9 grams); in efficacy studies, it was 51 grams (95% CI, −77 to −25 grams).

These findings suggest that brief counseling interventions in primary care settings lower alcohol consumption in those with nondependent unhealthy alcohol consumption. However, these effects are modest, and evidence for reduced consequences is not robust. This systematic review extends what we have known from other evidence reviews by specifically examining whether effects were similar in studies that assessed efficacy and effectiveness. Results were similar in both types of studies suggesting that brief counseling is effective in real world primary care settings.

However, the exclusion criteria for the original individual randomized trials (for example, alcohol dependence and psychiatric co-morbidity) limit the conclusions that can be drawn from this study. Nonetheless, the results do support practice guidelines that have recommended universal screening for unhealthy alcohol use in primary care settings and brief interventions for those who screen positive. Further research is needed in patient populations for whom the practice's benefits are less certain (e.g., medical inpatients, patients with dependence); research is also needed to determine how to go beyond decreasing consumption to decreasing consequences.

Intensive Referral to 12-Step Self-help Groups and 6-month Substance Use Disorder Outcomes[8]

Primary care clinicians commonly refer their patients with substance dependence to 12-step groups like Alcoholics Anonymous (AA) or Narcotics Anonymous (NA).[9] But making these referrals is often not simple; some patients may not be ready to take a step towards abstinence, and others have myriad ideological and practical barriers that prevent 12-step meeting attendance and participation. To examine whether intensive referral to AA or NA confers more benefit than standard, ad hoc referral, researchers randomized 345 veterans who entered a new substance abuse treatment episode (80% misused alcohol).[8] For the 181 patients assigned to intensive referral, treatment counselors provided a list of local AA/NA meetings preferred by other patients with directions to and times of those meetings and an introductory handout on 12-step groups. Counselors also arranged for an AA/NA volunteer to escort the patient to a meeting, confirmed with the patient in writing the meetings that they will attend, followed up on meeting attendance at the next session, and recommended in writing that the patient obtain a temporary sponsor. The 164 patients assigned to standard referral received only a schedule of local 12-step meetings and scripted encouragement to attend.

At six-month follow-up, intensive and standard referral groups did not differ on 12-step attendance overall, but intensive referral yielded more attendance among patients with less previous 12-step experience. The intensive group had more 12-step involvement (e.g., more likely to have participated actively in the meetings, had a spiritual awakening, or obtained a 12-step sponsor). The intensive group patients also improved more than the standard group on the Addiction Severity Index Alcohol Use (mean improvement 0.215) and Drug Use Composites scores (mean improvement 0.079). Twelve-step involvement mediated these effects. Finally, the intensive group patients were more likely to be abstinent from other drugs than the standard group (78% vs. 70%) and were more likely to be abstinence from alcohol, a difference of borderline significance (64% vs. 55%, p=0.06).

Although the referrals in this study were done by counselors among patients entering addiction treatment, the results have important implications for primary care clinicians. Internists' efforts to encourage substance-abusing patients to participate actively in AA/NA are likely worthwhile, especially for patients with less prior 12-step experience. Arranging for an AA/NA volunteer to escort the patient to a meeting and following up on participation appears to enhance successful facilitation of 12-step involvement. These findings are consistent with work from Project Match and others that suggest formalized 12-step facilitation is an effective relapse prevention strategy.[10–13] Furthermore, with modest effort and resources, the intensive referral implemented in this study could be reproduced in primary care settings, improving the management of substance dependence.

Topiramate for Treating Alcohol Dependence: A Randomized Controlled Trial[14]

Pharmacotherapy for alcohol dependence is approved by the FDA, but it is underutilized.

Topiramate may decrease alcohol consumption among alcohol-dependent persons by reducing dopamine release (and therefore alcohol's rewarding effects) via facilitation of GABA activity and inhibition of glutamate function, but has not been tested in a large controlled study. Investigators conducted a randomized trial of 371 alcohol-dependent men and women from 17 sites in the United States to determine if topiramate is more efficacious than placebo for reducing drinking.

Trial participants were alcohol-dependent adults (men who drank 35 or more drinks per week and women who drank 28 or more drinks per week; all participants scored 8 or higher on the Alcohol Use Disorders Identification Test (AUDIT).[15,16] To be enrolled, participants had to express a desire to stop or reduce alcohol consumption and be free of comorbid conditions. After extensive screening, trial participants were randomized to receive topiramate or placebo for 14 weeks, and both participants and investigators were blinded to treatment assignment. Medication was titrated during the first six weeks in scheduled increments to achieve a minimum topiramate (or placebo equivalent) dose of up to 50 milligrams per day and a maximum of 300 milligrams per day. All participants received weekly brief manual-guided adherence enhancement counseling.

In analysis, the researchers employed a conservative approach that assumed all dropouts to have relapsed to baseline drinking behaviors. With this consideration, topiramate recipients showed a greater reduction in the percent of drinking days than placebo recipients (from a mean of 82% to 44% vs. 82% to 52%, p=0.002), a greater increase in abstinent days than placebo recipients (from a mean of 9.7% to 37.6% vs. 9.4% to 29.1%, p=0.002), and a greater reduction in liver enzymes. Using a less conservative approach that considered dropouts as missing rather than relapsed, topiramate recipients showed even greater reductions in percent of drinking days than placebo recipients (from a mean of 82% to 20% versus 82% to 42%, p<0.001). With both analytic approaches, topiramate recipients achieved at least 28 days of both continuous abstinence and continuous non-heavy drinking faster than placebo recipients.

These results suggest that topiramate is a promising treatment for alcohol dependence for those seeking help with their drinking. The conservative analytic approach suggests that broadening the use of topiramate to treat alcohol dependence among adults who desire to reduce their alcohol consumption is warranted. Unfortunately, the side effects of topiramate (including depression, insomnia, difficulty with memory, somnolence, paresthesia, psychomotor slowing, dizziness, and nausea) may limit widespread acceptance as pharmacotherapy for alcohol dependence. Furthermore, because this randomized controlled trial had strict eligibility criteria to ensure that safety and efficacy could be measured, the generalizability of these findings to patients with other comorbid illnesses, such as other substance disorders or psychiatric disease, may be limited. In summary, this study, along with additional analyses that showed improvement in physical health and quality of life of patients on topiramate,[17] indicates that topiramate may be another pharmacotherapy available to physicians to treat alcohol dependence. Although the study is convincing regarding topiramate's positive effects on alcohol dependence, it is not FDA-approved for this indication. But with four approved drugs available (counting both injectable and oral naltrexone), prescription of pharmacotherapy represents an obvious means for primary care clinicians to become involved in the management of patients with alcohol dependence that is similar to how physicians address other chronic conditions like hypertension, diabetes and asthma.

OPIOID USE AND DEPENDENCE

Systemic Review: Opioid Treatment for Chronic Back Pain: Prevalence, Efficacy, and Association with Addiction[18]

Researchers systematically reviewed the literature to determine the prevalence and efficacy of opioid treatment for chronic back pain. They also examined the risk for substance use disorders and prescription medication misuse among patients perpetually treated with opioids.

Across 11 studies, the prevalence of opioid prescribing for chronic back pain ranged widely, from 3% to 66%. Regarding efficacy, in a meta-analysis of data from five studies, pain decreased (though non-significantly) with opioid treatment. Opioids had more efficacy than non-opioids or placebo in four of six short-term (less than four months) treatment studies. In four studies, the prevalence of a current substance use disorder in patients receiving opioids for chronic back pain also ranged widely (3–43%). These studies generally were of poor quality. In the highest quality study, the prevalence was 23%. It is notable that this prevalence (23%) was the same as in a comparison group of patients with chronic back pain who had not received opioid treatment. Across five studies, the prevalence of substance use disorders in patients receiving opioids for chronic back pain was 5% to 24%. These studies generally did not consider whether inadequate pain relief led to a misdiagnosis of substance misuse (known as pseudoaddiction[19]).

Although the quality of studies to date is limited and efficacy testing for chronic back pain has not been extensive, this collection of evidence suggests that opioids are a reasonable short-term treatment option. Nonetheless, the evidence for their efficacy is scanty, and long-term benefits are unknown. Furthermore, while the prevalence of a substance use disorder may or may not be higher than in other patients, these disorders are common particularly in people with chronic pain. It is critical to know whether such disorders are present when prescribing opioids for chronic pain so that they can be addressed.

Prescription Opioid Use, Misuse, and Diversion Among Street Drug Users in New York City[20]

To determine the patterns of prescription opioid use, misuse, and diversion among 586 drug users in New York City, researchers conducted detailed interviews. Among their findings, 72% of subjects used methadone and 65% sold it. Methadone was used and sold by more individuals than were long-acting preparations of oxycodone, hydrocodone/acetaminophen, or oxycodone/acetaminophen. More than half (58%) of prescription drug users reported that they used the opioids they obtained from physicians' offices to treat pain, prevent withdrawal, or to obtain euphoria. For example,

among the subjects who reported obtaining long-acting preparations of oxycodone from physicians, 83% reported using the medication primarily to treat pain, 50% primarily to prevent opioid withdrawal symptoms, and 38% primarily to experience euphoria. Of note, this study found that prescription drug users reported they were less likely to obtain prescription opioids for euphoria than for pain, and when they obtained prescription opioids for euphoria, they usually did so from dealers instead of physicians.

The primary limitations of this study as they relate to primary care are that the patients were not selected from clinical sites, but they did report on what they did with medications obtained from physicians. In addition, prescription drug use patterns are likely to vary geographically limiting the generalizability of these findings.

This study highlights a growing problem: abuse of and dependence on prescription opioids. Practicing physicians confronted with decisions about prescribing opioid medications need to weigh perceived benefits with potential adverse effects. In this study, methadone was the most commonly diverted prescription opioid, and many individuals used these medications to avoid opioid withdrawal or to treat pain; both findings are informative. Patients were less likely to use physician-obtained medications for euphoria than for other indications. Other work has shown that patient factors such as panic disorder, social phobia and agoraphobia, low self-rated health status, and other substance misuse among those with non-medical use of prescription opioids should alert clinicians to screen for prescription opioid abuse and dependence.[21] Regardless, this study highlights that physicians should prescribe opioids with caution and consider offering office-based treatment or specialty treatment referral when indicated.

Mortality Prior to, During and after Opioid Maintenance Treatment (OMT): A National Prospective Cross-Registry Study[22]

Opioid dependence is associated with significant morbidity and mortality.[23] Overdose is among the leading causes of death. This study, conducted in Norway, sought to determine the extent to which opioid agonist treatment reduced mortality in patients with opioid dependence. Researchers linked data collected over a seven-year period from a national death registry to a national database of people who were on a waiting list for opioid agonist treatment, receiving opioid agonist treatment (predominantly methadone), or who had discontinued opioid agonist treatment. Over the course of the study, 213 of 3789 patients died, 113 (53%) from overdose. Overdose mortality rates per 100 person years were 1.9 (95% CI, 1.6–2.1) for those on the waiting list for treatment, 0.4 (95% CI, 0–0.8) during opioid agonist treatment, and 2.1 (95% CI, 1.7–2.5) after treatment was discontinued. Overall mortality (relative risk [RR] 0.5; $p<0.001$) and overdose mortality (RR 0.2; $p<0.001$) were lower in patients receiving opioid agonist treatment than in patients on the waiting list. Among those who discontinued treatment, total mortality risk was higher among men (compared with men on the waiting list) (RR, 1.8; $p<0.02$), but not among women.

The primary limitation of this study relates to the regulations regarding entry into opioid agonist treatment in Norway, which are more stringent than those in the United States. The Norwegian regulations restrict this treatment to individuals who are at least 25 years of age and can demonstrate several years of opioid dependence. Patients with medical and psychiatric co-morbidity are given priority access to these services.

This investigation adds to the ample evidence that opioid agonist treatment reduces mortality in opioid-dependent patients. In the arena of pain treatment, cases of overdose death that appear to be attributed to physician-prescribed methadone have increased the potential for negative public and regulatory backlash against methadone.[24,25] Therefore, these results may play an important role in contemporary policy discussions about opioid agonist treatment for opioid dependence. Furthermore, primary care clinicians can use this information to decide to provide or advocate for access to initial and maintenance opioid agonist treatment (e.g. with buprenorphine) for their patients with opioid dependence.

Treating Homeless Opioid Dependent Patients with Buprenorphine in an Office-based Setting[26]

Buprenorphine treatment outcomes are generally evaluated in resource-rich settings (e.g., with research staff) or among patients with some social support. The effectiveness of this treatment in everyday practice settings and among indigent patients remains unclear. Two studies explored more generalizable approaches to buprenorphine treatment for opioid dependence. In one study, a Boston group compared the effectiveness of buprenorphine in patients treated at a clinic for the homeless (n=44) and in-house patients treated in a general primary care setting (n=41). A nurse care manager was actively engaged in patients' care at both sites. Although homeless patients were more likely than housed patients to have comorbidity, treatment outcomes were similar between the two groups. Twenty-one percent of homeless patients and 22% of housed patients "failed treatment" (were lost to follow-up during induction phase or discharged due to disruptive behavior or ongoing alcohol or other drug use while not adhering to intensified substance abuse treatment). Both groups had median treatment retention of nine months and equally low rates of illicit opioid use at 12 months. Homelessness resolved for 36% and employment rates increased in both groups.[26]

This study supports the premise that buprenorphine opioid agonist therapy can be applied to diverse patients in typical healthcare environments. These findings are corroborated by other investigators who examined 99 patients receiving buprenorphine treatment in 1) a hospital-based primary care center with an on-site pharmacy but no on-site addiction counselor or 2) a neighborhood health center with an on-site addiction counselor but no on-site pharmacy.[27] At six months, 54% of patients were abstinent from illicit opioids (determined by urine toxicology, self-reported drug use, and general clinical assessment). Clinical outcomes did not differ across the two treatment settings.

The findings of both feasibility studies support the effectiveness of extending office-based buprenorphine treatment into less specialized, low-intensity settings and to patients with only marginal social support. The results imply that these interventions can be delivered in real world primary care settings. Implementation and application of evidence based

addiction treatment, such as buprenorphine, into typical clinical settings may reduce addiction- and non-addiction-related morbidity and improve the quality of care provided to patients with addictions.

Acknowledgements: *All authors are members of the Society of General Internal Medicine's Substance Abuse Interest Group to which we are deeply indebted for its support. The authors also wish to thank Margaret Krumm for her contributions of coordinating the review and editing this paper.*

Conflict of Interest: *No author has significant conflict of interest that directly applies to this manuscript. However, Dr. Friedmann has received honoraria (Reckitt-Benckiser), Dr. Gourevitch has received a grant from a for-profit entity (Alkermes, Inc.), Dr. Arnsten has received consultancy reimbursement (Proteus Biomedical, Inc), and Dr. Saitz has received reimbursement for expert testimony.*

Corresponding Author: *Adam J. Gordon; VA Pittsburgh Healthcare System, University of Pittsburgh School of Medicine, Mailcode 151-C-H, 7180 Highland Drive (151-C-H), Pittsburgh, PA 15206, USA (e-mail: adam.gordon@va.gov).*

REFERENCES

1. **Whitlock EP, Polen MR, Green CA, Orleans T, Klein J.** 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. Ann Intern Med. 2004;140(7):557–68.
2. **Mokdad AH, Marks JS, Stroup DF, Gerberding JL.** Actual causes of death in the United States, 2000. JAMA. 2004;291(10):1238–45.
3. Alcohol, other drugs, and health: current evidence. Available at: http://www.bu.edu/aodhealth/about.html. Accessed June 1, 2008.
4. **Gordon AJ, Sullivan LE, Alford DP, et al.** Update in addiction medicine for the generalist. J Gen Intern Med. 2007;22(8):1190–94.
5. **Vinson DC, Kruse RL, Seale JP.** Simplifying alcohol assessment: two questions to identify alcohol use disorders. Alcohol Clin Exp Res. 2007;31(8):1392–8.
6. **Kaner EF, Beyer F, Dickinson HO, et al.** Effectiveness of brief alcohol interventions in primary care populations. Cochrane Database Syst Rev. 2007;(2):CD004148.
7. National Institute on Alcohol Abuse and Alcoholism. Helping patients who drink too much: a clinician's guide (NIH Publication No. 05–3769). Rockville, MD: U.S. Department of Health & Human Services; 2005.
8. **Timko C, DeBenedetti A, Billow R.** Intensive referral to 12-Step self-help groups and 6-month substance use disorder outcomes. Addiction. 2006;101(5):678–88.
9. **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(2):84–91.
10. Project MATCH Research Group. Matching alcoholism treatments to client heterogeneity: treatment main effects and matching effects on drinking during treatment. J Stud Alcohol. 1998;59(6):631–9.
11. Project MATCH secondary a priori hypotheses. Project MATCH Research Group. Addiction. 1997;92(12):1671–98.
12. Project MATCH Research Group. Matching alcoholism treatments to client heterogeneity: Project MATCH three-year drinking outcomes. Alcohol Clin Exp Res. 1998;22:1300–11.
13. **Timko C, DeBenedetti A.** A randomized controlled trial of intensive referral to 12-step self-help groups: one-year outcomes. Drug Alcohol Depend. 2007;90(2–3):270–9.
14. **Johnson BA, Rosenthal N, Capece JA, et al.** Topiramate for treating alcohol dependence: a randomized controlled trial. JAMA. 2007;298(14):1641–51.
15. **Conigrave KM, Hall WD, Saunders JB.** The AUDIT questionnaire: choosing a cut-off score. Alcohol Use Disorder Identification Test. Addiction. 1995;90(10):1349–56.
16. **Allen JP, Litten RZ, Fertig JB, Babor T.** A review of research on the Alcohol Use Disorders Identification Test (AUDIT). Alcohol Clin Exp Res. 1997;21(4):613–9.
17. **Johnson BA, Rosenthal N, Capece JA, et al.** Improvement of physical health and quality of life of alcohol-dependent individuals with topiramate treatment: US multisite randomized controlled trial. Arch Intern Med. 2008;168(11):1188–99.
18. **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.
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–S60.
20. **Davis WR, Johnson BD.** Prescription opioid use, misuse, and diversion among street drug users in New York City. Drug Alcohol Depend. 2008;92(1–3):267–76.
21. **Becker WC, Sullivan LE, Tetrault JM, Desai RA, Fiellin DA.** Non-medical use, abuse and dependence on prescription opioids among U.S. adults: psychiatric, medical and substance use correlates. Drug Alcohol Depend. 2008;94(1–3):38–47.
22. **Clausen T, Anchersen K, Waal H.** Mortality prior to, during and after opioid maintenance treatment (OMT): a national prospective cross-registry study. Drug Alcohol Depend. 2008;94(1–3):151–7.
23. **Hser YI, Hoffman V, Grella CE, Anglin MD.** A 33-year follow-up of narcotics addicts. Arch Gen Psychiatry. 2001;58(5):503–8.
24. **Graham NA, Merlo LJ, Goldberger BA, Gold MS.** Methadone- and heroin-related deaths in Florida. Am J Drug Alcohol Abuse. 2008;34(3):347–53.
25. Center for Substance Abuse Treatment. Summary report of the meeting: methadone mortality- a reassessment. Available at: http://www.dpt.samhsa.gov/pdf/Methadone_Report_10%2018%2007_Brief%20w%20attch.pdf. Accessed July 1, 2008.
26. **Alford DP, LaBelle CT, Richardson JM, et al.** Treating homeless opioid dependent patients with buprenorphine in an office-based setting. J Gen Intern Med. 2007;22(2):171–6.
27. **Mintzer IL, Eisenberg M, Terra M, et al.** Treating opioid addiction with buprenorphine-naloxone in community-based primary care settings. Ann Fam Med. 2007;5(2):146–50.

Published in final edited form as:
Adv Genet. 2008 ; 62: 33–65. doi:10.1016/S0065-2660(08)00602-0.

Genetics of the Framingham Heart Study Population

Diddahally R. Govindaraju[1], **L. Adrienne Cupples**[2], **William B. Kannel**[3], **Christopher J. O'Donnell**[3], **Larry D. Atwood**[1,2], **Ralph B. D'Agostino Sr.**[2,3], **Caroline S. Fox**[3], **Marty Larson**[3], **Daniel Levy**[3], **Joanne Morabito**[3], **Ramachandran S. Vasan**[3], **Greta Lee Splansky**[3], **Philip A. Wolf**[1], and **Emelia J. Benjamin**[3,5,6]

[1]Department of Neurology, Boston University School of Medicine, Boston, USA

[2]Department of Biostatistics, Boston University School of Public Health

[3]NHLBI's Framingham Heart Study

[4]Department of Neurology, Boston University School of Medicine

[5]Department of Cardiology, Boston University School of Medicine

[6]Department of epidemiology, Boston University School of Public Health

Abstract

This article provides an introduction to the Framingham Heart Study (FHS) and the genetic research related to cardiovascular diseases conducted in this unique population1. It briefly describes the origins of the study, the risk factors that contribute to heart disease and the approaches taken to discover the genetic basis of some of these risk factors. The genetic architecture of several biological risk factors has been explained using family studies, segregation analysis, heritability, phenotypic and genetic correlations. Many quantitative trait loci underlying cardiovascular diseases have been discovered using different molecular markers. Additionally, results from genome-wide association studies using 100,000 markers, and the prospects of using 550,000 markers for association studies are presented. Finally, the use of this unique sample in genotype and environment interaction is described.

"Nature is all that a man brings with himself into the world; nurture is every influence from without that affects him after his birth."

- Francis Galton (1890, p. 9)

"Why should you,…put yourself to the trouble of being measured, weighed and otherwise tested? Why should I…and why should others, take the trouble of persuading you to go through the process?…A comparison of the measures made from time to time will show whether the child maintains his former rank, or whether he is gaining on it or losing it."

- Francis Galton (1890)

I. INTRODUCTION

Coronary heart disease (CHD) has remained a major cause of morbidity and mortality in the United States, affecting nearly 13 million people and causing approximately one million deaths per year (Thom *et al.*, 2006). Although the incidence of cardiovascular diseases (CVDs) has gradually declined since the 1960s in the U.S. (Cooper *et al.*, 2004), it is

---

The terms, Framingham Heart Study population and Framingham Heart Study cohort are used interchangeably.

reaching epidemic proportions in many countries of Europe and the developing world (Yusuf *et al.*, 2001). In the 1940s CHD was recognized as the leading cause of mortality in the U.S. accounting for approximately half of all deaths (Kannel, 1990). Nonetheless, knowledge of the factors that disposed individuals to CVDs was "virtually non existent" 60 years ago and was perceived to be an inevitable consequence of "aging or genetic predisposition" of individuals (Dawber and Kannel, 1999). Fortunately, the U.S. Public Health Service (USPHS then and later NIH) recognized the necessity for understanding the causal factors of the epidemic and decided to establish a prospective longitudinal observational epidemiological study in 1947, in the town of Framingham, Massachusetts in collaboration with the Massachusetts State Department of Health and Harvard Medical School. The "Framingham Study" was formally established in 1948, to identify factors that contribute to CVD (Dawber *et al.*, 1951; Kagan *et al.*, 1962; Levy and Brink, 2005).

The study, nearly six decades later and now known as the "Framingham Heart Study" (FHS), is the longest running, multigenerational longitudinal study in medical history (Butler, 1999). It has helped identify several "risk factors" and their cumulative influence on the manifestation of CVD. Indeed, the term 'risk factor' was coined by Framingham investigators (Kannel *et al.*, 1961).

Information collected on the participants enrolled in the study has aided in correcting a number of long held misconceptions on the role of blood pressure, lipids, diabetes, obesity, proteinuria, left ventricular hypertrophy, atrial fibrillation, smoking and exercise in the manifestation of CVD. Framingham investigators have also elucidated the pathogenesis of atherosclerosis and thus have laid a firm foundation toward preventive cardiology (Kannel, 1990). Furthermore, the study has acquired an iconic status in public health and preventive cardiology and has been listed as the "fourth significant achievement in medicine" (after the development of antibiotic treatments, immunization against infectious diseases, and the understanding of the roles of vitamins; Anon., 1999), and the second greatest discovery (behind electrocardiography) in Cardiology (Mehta and Khan, 2002).

The investigators of the *original protocol* of the "Framingham Study" recognized a wide range of variation among individuals in human populations in response to "stresses and insults" (Gordon and Kannel, 1970). Instead of focusing on just one or a few independent causal factors that might influence CVD, they took an integrated approach and hypothesized that CVD may arise from "multiple causes which work slowly within the individual." However, family history for CVD received the highest importance among many variables selected for studying its manifestation among the participants (Dawber *et al.*, 1951). In general, at least three major variables were assumed to contribute to the onset of CVD: constitutional (heredity), and conditioning (environmental) factors, as well as the length of time taken by the conditional factors to act on constitutional factors ultimately resulting in a clinically recognizable condition (Gordon and Kennel, 1970). Thus, the founders of the study were cognizant of the fact that the biological basis of CVD may be complex and may be modulated by the interaction of heredity and environmental factors.

Although the role of hereditary factors in the development of CVD was acknowledged from the very beginning of the Framingham study, genetic studies did not receive much attention until the late 1980s. In the last twenty years, however, a number of investigators have utilized the rich resource available at the study and have attempted to understand the genetic basis of CVD using various approaches. In this review, we briefly discuss: a) some of the salient features of the Framingham Heart Study population, and b) approaches taken by the Framingham investigators toward identifying the genetic bases of CVD and some of its risk factors.

II. The study population

A. Demography

The Framingham Heart study is comprised largely of whites of European descent. Individuals from the Italian, Irish and English ancestry are predominant in the sample. About 85 percent of the Original Cohort was born in the U.S. or Canada, including 19% born in Framingham and another 40% born in Massachusetts. Thirty-five percent identify themselves with ethnic origins in the British Isles, including 15% from Ireland; another 19% are of Italian ethnicity, 32% of other Western European ancestry, 5% Canadian and 6% Eastern European. Less than 4% are of non-European origins or of unknown ethnicity (Table 1).

B. Multigenerational cohorts and examinations

The study was formally established from 1948–1953 in the town of Framingham MA, located about 20 miles west of Boston. Approximately 10,000 individuals were found to be of ages 30–59 years from a total population of 28,000. It was determined that if 6,000 individuals were invited into the study from the 10,000 in the target age range, about 5,000 individuals would not only be free of cardiovascular disease, but also provide sufficient sample size for the analysis of factors contributing to the development of CHD among the selected individuals over a period of twenty years. In such a time span, approximately 400, 900, and 2,150 would develop CHD at the end of 5^{th}, 10^{th} and 20^{th} year, respectively, from the initial examination. Therefore, all the households in Framingham listed in the town census were categorized by the number of eligible individuals, and every third household was excluded. Approximately 6,600 individuals were so selected. As expected the number was diminished by losses, deaths and refusals. There were also 740 volunteers from the town of Framingham included. At the beginning of the study, 5,209 men and women joined from January 1948 through March 1953 (Kagan *et al.*, 1962), and a total of 5,128 these participants were found to be free of “overt coronary heart disease”. Thus, the group consisting of 5,209 participants constitutes the “Original cohort” of the study. The participants would undergo examinations every two years (Dawber *et al.*, 1951). The Offspring and Third Generation cohorts consisting of 5124 and 4095 individuals, respectively, were recruited in 1971–1975 and 2002–2005, respectively. The Offspring cohort was comprised of children of the Original cohort and the children’s spouses and has been examined every 4 to 8 years. The Third generation cohort was recruited from children of the Offspring. The participants of the Original, Offspring, and Third Generation cohorts have been examined 29, 5 and one time, respectively, for a large number of variables that may have a bearing on CVDs. (Figure 1, 2). Thus, most participants of the study are members of 754 extended pedigrees. These pedigrees are well-defined (parents, children, grand children, cousins, avuncular cousins, aunts, uncles etc.) and range in size from 3 to 230 individuals with a median of 9 (Figure 3); the third generation individuals represent 1828 nuclear families whose sibships vary from 1 to 9 individuals.

C. Diversity of traits measured – Phenotypic, physiological and environmental

At the initiation of the study, a committee consisting of eleven physician-epidemiologists developed a list of criteria and measured variables that may have a “bearing on the development of (CV) disease” under the following six categories (see Dawber *et al.*, 1951 for details).

- **i.** Medical history - family history of CVD among parents, siblings and children, symptoms such as chest pain, sleeping habits, alcohol and tobacco consumption
- **ii.** Physical examination - aimed at detecting cardiovascular abnormalities and diseases as well as height, weight, chest and waist circumference, thyroid

enlargement, pulmonary disease, cardiac murmurs or gallops, blood pressure, liver enlargement, varicose veins

iii. Chest X-ray examination

iv. Electrocardiogram using 12 leads. Electrokymographic tracing at 12 points on the cardiac silhouette

v. Blood examination for hemoglobin, serum cholesterol, phospholipids, glucose concentrations

vi. Urine analysis

This tradition of routine ascertainment of physical examination, life style and habits, medical history, laboratory analysis, non-invasive and end-point data has been applied to all three generations of participants (Table 2). The number of variables, however, has increased over time and varies from one exam to the next. For example, in the first examination, data were collected on 30 major variables. Over time, the diversity and complexity of phenotyping has expanded. For instance, in recent examinations the Offspring, aside from standard history and physical examination measures, have undergone additional testing including carotid ultrasound, echocardiography, brachial reactivity, arterial tonometry, 6 minute walk, pulmonary function testing, and subsets have received cardiac and brain magnetic resonance imaging, cardiac multidetector computed tomography, and bone densitometry (www.nhlbi.nih.gov/resources/deca/fhsc/docindex.htm). In a recent survey, approximately 1500 variables were found to have been measured on the FHS cohort. However, not all traits have been measured on all of the individuals; hence, the number of phenotypes measured varies among individuals, examination cycles and cohorts.

D. Multifactorial nature of the heart disease

Cardiovascular diseases arise from multiple causes. The heterogeneous nature of the etiology of CVD was recognized at the start of the study in 1948. Several key factors either independently or cumulatively were found to exert influence disproportionately to the development of CVD. These factors were designated as “risk factors.” The primary risk factors include: age, systolic blood pressure (SBP), body mass index (BMI), total/HDL ratio, diabetes, and smoking (Dawber *et al.*, 1959; Kannel *et al.*, 1961). Additional risk factors and their components including morphological (e.g. left ventricular hypertrophy; Kannel *et al.*, 1969), physiological (e.g., fibrinogen; Kannel *et al.*, 1987), and life style (e.g. Posner *et al.*, 1993) have been added over time. These are further categorized into modifiable, probably modifiable, and fixed risk factors (Table 3; Wilson, 1994). Distribution of various risk factors in all the three cohorts as well as between men and women is provided in Table 4.

III. Phenotypic and genetic architecture of complex traits

Biological variation may be understood at two levels: phenotypic and genetic. Many of the CVD risk factors such as high density lipoprotein cholesterol, total cholesterol, and blood pressure are quantitative traits. The phenotypic variation of a quantitative trait may be represented by $V_P = V_G + V_E + 2cov_{GE}$, where G, E and $2cov_{GE}$ are genetic, environmental and their interaction variances, respectively (Falconer and Mackay, 1996). An understanding of the genetic architecture of a quantitative trait requires knowledge of its inheritance pattern, association with other traits and molecular characterization of genes that underlie the phenotype (Mackay and Lyman, 2005). Complex diseases such as CVD may arise from multiple genes and their interaction with environmental factors. Hence, it is important to tease apart the components that contribute toward the development of these diseases using genetic approaches. The Framingham Heart Study provides a unique opportunity for understanding the genetic architecture of many human traits, including the CVD risk factors,

using the detailed family structure, detailed phenotypic measurements, information on physiological and molecular markers. Although the original protocol of the FHS recognized the role of heredity and environmental factors in CVD, systematic genetic analysis did not start until the mid 1980s. DNA collection on each of the participants from the Original and Offspring cohorts was initiated in the late 1980s, continued during the 1990s and was expanded to Third Generation participants at their first examination.

A. Inheritance patterns of CHD

1. Family studies—The fact that both morphological and disease traits cluster in families has been known to human geneticists for a long time (Galton, 1886; Garrod, 1902), and family history is a significant predictor of heart diseases (Friedlander et al 1985). The Framingham investigators indeed recognized the fact that CHD "runs in families" (Kannel *et al.*, 1979; Kannel and Stokes, 1985); yet the relative contribution of genetic factors and shared environment toward developing cardiovascular risk was debated, since "family members eat at the same table" (Kannel *et al.*, 1979; Kannel and Stokes, 1985). On the contrary, Havlik *et al.*, (1979) reported significant correlations between parents and offspring and sibling pairs for blood pressure. Correlation between spouses was attributed to assortative marriages for age, body weight and habits such as smoking and alcohol consumption. Similarly, Myers *et al.*, (1990) demonstrated that CVD in parents could be an independent risk factor. Similar studies have been carried out at the FHS for other traits such as cardiac heart disease (Brand *et al.*, 1992), lens opacities (Anon., 1994), stroke and hypertension (Reed *et al.*, 2000), atrial fibrillation (Fox *et al.*, 2004) and heart failure (Lee *et al.*, 2006). Many of the risk factors, discovered by the FHS investigators, act cumulatively toward determining CVD risk between parents and offspring (Figure 4; Lloyd-Jones *et al.*, 2004).

Family studies point toward the aggregation and inheritance of disease causing factors among individuals within families. They do not, however, indicate if the mode of genetic transmission from parent to offspring is simple or complex. Segregation analysis, on the other hand, provides insights on whether or not the inheritance is Mendelian (simple) or complex. For example, using the FHS family data, Felson *et al.*, (1998) reported the presence of a major recessive gene and a multifactorial component for generalized arthritis. On the other hand, pulmonary function was found to be governed by a polygenic component (Givelber *et al.*, 1998). Interestingly, a number of risk factors appear to differ among men and women (Table 4), which could ultimately contribute to their susceptibility to CVD (Figure 5; Hubert *et al.*, 1983).

2. Heritability and Genetic correlations—The relative contribution of genetic and environmental factors on the expression of quantitative traits is determined using the index known as heritability. Formally, heritability represents the amount of phenotypic variability or variance explained by genetic factors and is estimated as a ratio of genetic to phenotypic variance. Either broad (H^2) or narrow sense (h^2) estimates are used for this purpose (Sham, 1988). By definition, broad sense heritability includes all genetic variance (both additive, dominance and their interaction), but the narrow sense heritability considers only the additive portion of the genetic variance (Falconer and Mackay, 1996). Heritability serves two purposes: it provides an estimate of the level of genetic variation underlying a quantitative trait, including disease, and also indicates the evolutionary potential of the trait (Lynch and Walsh, 1998). In general, moderate to high heritability has been reported for most traits examined (Table 5), but the distribution of heritability for the many traits examined in this highly phenotyped cohort is unknown. Note that heritability is a population estimate, and therefore, it could vary across populations, between sexes, environments as well as at different stages in the life span (Lynch and Walsh, 1998). These instabilities of

heritability estimates are also seen for various traits in the FHS sample (Table 5). For example Brown *et al.*, (2003) demonstrated a general decrease in estimated heritability in 70 versus 40 year old individuals (Figure 6). Furthermore, Atwood *et al.*, (2005) indicated that heritability for white matter hypersensitivity decreased in women, but increased slightly in men with advancing age (Figure 7).

A number of morphological and biochemical traits are correlated, and these associations that may be attributed to three factors: genetic, developmental and environmental (Lynch and Walsh, 1998). Thus, any variation in the relations among traits, either due to environmental or age-related changes, may reflect the effects of underlying genes and common genetic precursors, developmental pathways as well as coordinated organism wide-signaling (Badyayev and Fresman, 2004). Genetic correlations among traits arise from pleiotropic effects of genes on multiple traits and/or linkage disequilibria among distinct loci (Cheverud, 2001). Genetic correlations could also reflect allelic complexes at multiple loci as well as coadaptation (Churchill, 2006). Conversely, genetic correlations might indicate widespread association among loci, due to linkage and/or pleiotropy at the genomic level, which in turn could govern the integration of both morphological traits and disease related traits (Churchill *op cit.*,). Phenotypic, genetic and environmental correlations have been determined among five risk factors (cholesterol, high density lipoprotein, systolic blood pressure, triglycerides and body mass index) in the FHS (Table 6). The results indicate that the phenotypic and genetic correlations have similar magnitudes. In other cases, whereas the magnitude differed, the direction of the correlation was conserved. Additionally, the concentrations of high density lipoprotein and triglycerides were affected by environment. These results largely agree with the conclusions reached by Cheverud (1988), who suggested that phenotypic correlations may reflect genetic correlations.

3. Physiological and molecular markers—Phenotypes are linked to genes via biochemical pathways, and therefore, biochemical (bio) markers or biological traits provide logical surrogates to establish the relations between disease phenotypes and genotypes. These molecules or traits, also called endophenotypes or risk factors, in turn reflect the action of underlying genes and their expression patterns (Rice et al. 2001). Hence, measuring informative biochemical markers to predict the behavior of phenotypes is often favored in CVD (Vasan, 2006), as they simultaneously provide an idea of the phenotypes, genes and the pathway. A number of biomarkers have been used to establish relations between biomarkers and risk for cardiovascular disease in the FHS population. For example, Seman *et al.*, (1999) reported a positive association between lipoprotein (a) cholesterol concentrations and CHD in men but not in women. Keaney *et al.*, (2004) determined that ICAM-1 concentrations were associated with age, female gender, total/high density cholesterol ratio, body mass index, blood glucose, smoking and prevalent CVD. Similarly, Wang *et al.*, (2002) reported a close association between the concentrations of C-reactive protein, and carotid atherosclerosis, but the relationship was found only in women and not in men. High concentrations of total homocysteine have been implicated in cardiovascular disease (Arnesen *et al.*, 1995) and dementia (Seshadri et al. 2002). Elias *et al.* (2005) reported an inverse relation between the concentrations of homocysteine and cognitive function, only among individuals over 60 years in the FHS population.

In humans, a number of other classes of molecular markers have been employed to describe both genetic variation and to discover the genetic basis of phenotypic traits including complex diseases. These include: allozymes (Harris, 1966), restriction fragment length polymorphisms (RFLPS; Solomon and Bodmer 1979; Botstein *et al.*, 1980), variable number of tandem repeats (VNTRs; Jeffreys et al. 1985), and microsatellites (Weber and May, 1989) and more recently, single nucleotide polymorphisms (SNPs). Briefly, RFLPs are the products obtained by digesting the DNA molecules with restriction enzymes;

microsatellites are two [e.g. (CA)n] to five [(TTTTA)n] repeat sequences found distributed throughout the genome and are known to be highly polymorphic. SNPs arise from mutations at specific nucleotides in the DNA molecule and represent the most abundant class of polymorphisms in the human genome (see Strachan and Read 2003, for details).

The Framingham investigators have utilized primarily three families of molecular markers - RFLPs, microsatellites and SNPs - to establish associations between molecular markers and cardiovascular risk factors. For example, Fabsitz *et al.*, (1989) tested the association between human leukocyte antigen (HLA) and obesity on 348 individuals and found that the Bw35 allele was significantly associated with obesity. Similarly, RFLPs (for restriction enzymes, *Msp*I, *Pst*I, *Sst*I, *Pvu*II, *Xba*I) in the Aplolipoprotein gene cluster A-I, C-III, and A-IV were tested (Ordovas et al., 1991) on 202 patients with coronary artery disease and 145 normal individuals. They found that individuals with SstI had 38 percent greater concentration of triglycerides than the referents. Wilson *et al*, (1994), examined the relationship among the ε_2, ε_3 and ε_4 alleles of the apolipoprotein E locus in relation to CHD among 1034 men and 916 women aged 40 – 70. They found that ε_4 allele was associated with elevated low density lipoprotein cholesterol concentrations, as well as CHD in both men and women.

IV. Linkage and association studies

The availability of detailed measurements on cardiovascular risk factors and other phenotypic information in the FHS has facilitated mapping complex traits using two well known approaches: linkage and association. Briefly, linkage arises if two loci physically occur on the same chromosome and are inherited as a unit. It is determined using information on the inheritance pattern between parents and offspring in pedigrees (see Terwilliger and Ott, 1994 for details).

Linkage methods are used to identify regions at various locations on chromosomes or the genome that influence a given trait. These regions are assumed to contain quantitative trait loci (QTL).

Discovery of QTLs has been accomplished using primarily two types of linkage analyses: model based (parametric) and model free (non-parametric). In the former, a number of parameters such as the mode of inheritance of the disease, frequency of the causal allele, and its penetrance must be specified *a priori*. The likelihood of genetic linkage between two loci is determined by a LOD (logarithm of odd) scores. In general, for a Mendelian disorder, a LOD score of >3.0 is considered evidence for linkage (Sham, 1998). Parametric approaches have been successfully used for identifying the genetic basis of simple Mendelian disorders. Cardiovascular disorders reveal complex or non-Mendelian inheritance patterns that make it difficult to assign inheritance patterns. Therefore, model free analysis, which does not require *a priori* definition of allele frequencies or mode of inheritance, is used to map quantitative traits. This approach requires that the identity of specific alleles or set of linked alleles (haplotypes) that are inherited among relatives be identified, by means of identity-by-decent (IBD). In other words, IBD is central to model free linkage analysis. Model free approaches have been used at the FHS more extensively to understand the genetic bases of quantitative traits employing microsatellite markers.

A. Mapping with microsatellite markers

Approximately 612 microsatellite markers have been typed on the largest 330 pedigrees consisting of 1702 individuals belonging to generations 1 and 2 of the FHS. These data have been used to map genes underlying several risk factors, including blood pressure, arterial stiffness, lipid traits, adiposity glycemic traits, circulating biomarkers (e.g. inflammation,

natriuretic peptide), pulmonary function, renal function, and bone traits (Table 7). A number of these locations have been confirmed using other populations as replicate samples. Recently, the third generation individuals have been typed with a comparable set of STR markers. Upon completion, microsatellite markers will be available on about 7000 individuals, encompassing three generations, and linkage analyses will be extended to three generation pedigrees.

Besides identifying candidate loci for a number of risk factors, the availability of correlated traits and longitudinal data on families has facilitated FHS researchers to ask additional interesting questions. For example, does age variation influence the magnitude of LOD scores? Or does it lead to a shift in the location of a candidate gene region? Also, are several distinct yet correlated phenotypes influenced by the gene(s) located in a specific region? For instance, it is known that decreased high-density lipoproteins are inversely correlated with high cardiovascular risk. Arya *et al.*, (2003) mapped the region harboring genes that influence both BMI and HDL-C and thereby suggested pleiotropic effects. Similarly, Lin (2003) reported a common region, 6q24.3, to be influencing two inversely correlated traits, plasma triglycerides (PG) and high density lipoprotein cholesterol levels. Atwood *et al.*, (2006) on the other hand, performed linkage analysis on body mass index across 28 years to determine the impact of measurement across age groups. The results indicated that although the magnitude of LOD scores varied across six measurements ranging from 0.61 to 3.27, they all mapped to 11q14, suggesting that at least a QTL in this region for BMI may not be due to measurement errors.

B. Association studies

Linkage studies have been employed to map numerous genes underlying Mendelian diseases. This approach, however, is less powerful to map complex disorders as they are governed by many genes and their causal alleles whose effects are generally low. As noted earlier, parametric linkage approaches work best when the effect of the causal allele is large and least influenced by environmental factors. Complex traits, on the contrary, are greatly affected by environmental factors, making it more difficult to use linkage analysis. Risch and Marikangas (1996) proposed an alternative solution to this problem. They conjectured that association studies, using a large number of markers (in the neighborhood of a million) may be more useful for studying the genetic bases of complex disease than linkage studies. In association studies, a comparative analysis of alleles between individuals that carry the disease and healthy individuals is carried out, with the important assumption that the marker may be embedded in the causal gene or close to it. Additionally, association studies may or may not require pedigree information and could also be performed using samples that are unrelated or family-based. This approach has been feasible by the discovery and deployment of the most abundant class of molecular markers – single nucleotide polymorphisms (SNPs) – for association studies (see below).

Usually, two approaches are taken to establish an association between a putative causal site within a known gene (or any unknown site in the entire genome), with a given phenotype. Markers are placed at regular intervals along the length of the gene or across the genome, with the assumption that the markers so placed may be in linkage disequilibrium (LD) with the causal allele. In other words, information on how a marker can predict the presence or absence of disease causing alleles or locus could be determined using a linkage disequilibrium approach. Briefly, linkage disequilibrium is an index of non-random association of two alleles on a chromosome in a population (Ardlie *et al.*, 2002). If a new mutation occurs at any location of the genome, it is in complete linkage disequilibrium with the surrounding marker alleles. Among several measures proposed to measure linkage disequilibrium (Devlin and Risch, 1995), two methods, D' (Lewontin, 1964), and r^2 (Hill and Weir, 1994) are most frequently used. Accordingly, strong LD between the marker and

a causal allele (>0.8) is used as an index toward identifying a functional allele. Both of these two approaches have been used in FHS data and some of these results are presented below.

1. Association of known polymorphisms in candidate genes with cardiovascular risk factors—Causal polymorphisms within a number of candidate genes that affect the cardiovascular pathway have been described in the literature. FHS investigators have typed the same polymorphisms in FHS participants to confirm or refute the previously published associations. Examples include the association between two polymorphisms in the estrogen receptor-β gene with left ventricular mass and wall thickness in women with hypertension (Peter *et al.*, 2005); L162 polymorphisms of the peroxisome proliferator-activated receptor alpha (*PPARA*) and plasma lipids (Tai *et al.*, 2002); ATP-binding cassette transporter -1 (*ABCA1*; polymorphisms with HDL concentrations (Brousseau *et al.*, 2001). Additionally, SNPs in 200 genes of the cardiovascular pathway have been typed and a number of association studies have been performed with the following six echocardiographic phenotypes: left ventricular (LV) mass, LV internal dimension, LV wall thickness, left atrial dimension and aortic dimension and part of the results are presented in a grid form (http://cardiogenomics.med.harvard.edu/home; Levy *et al.*, 2006). Occasionally, however, a single SNP may suggest weak or no association with a given phenotype, but several SNPs in linkage disequilibrium (also known as haplotypes) may improve the strength of association. For example, Kathiresan *et al.*, (2006) found a triallelic haplotype containing C-T-A alleles of the C-reactive protein gene to be associated with serum C-reactive concentration.

2. Genome-wide association studies—Whereas linkage and candidate gene studies have revealed many potential regions and SNPs of interest, there have been relatively few successes in uncovering a comprehensive set of genetic variants responsible for common complex disease (Carlson et al. 2004). Meta-analyses of candidate gene studies suggest that only about 1/3 of the reported associations are validated, and less than 100 reported genetic associations are considered to be definitive (Lohmueller et al. 2003; Ioannidis et al. 2003). A limitation of candidate gene studies is that they are constrained by existing, often incomplete knowledge of the pathophysiology of disease. Technological breakthroughs in high throughput genotyping using 100 – 500 thousand well characterized, informative markers – single nucleotide polymorphisms (SNPs) – in combination with novel analytical techniques have opened the possibility of conducting genome-wide association studies. These approaches have also received an additional impetus from the success of the HapMap project (Altshuler *et al.* 2005; http://www.hapmap.org). The discovery and replication of the association between *CFH* (Complement Factor – H) gene and age-related macular degeneration, using informative SNPs obtained thorough the HapMap provided an early indication of the power of genome-wide association studies to accelerate gene discovery (Klein et al. 2005). The Framingham investigators have taken a two-tier approach to conduct genome-wide association studies using both 100,000 and 550,000 single nucleotide polymorphisms chips provided by Affymetrix.

a. 100K Study in the FHS population: In 2005 an Affymetrix 116K SNP genome-wide scan was conducted in about 1350 family members of the Original and Offspring cohorts of the FHS. Herbert and colleagues identified a common genetic variant associated with BMI near the *INSIG2* gene in Framingham participants; they replicated the finding in most of the other cohorts they tested (Herbert et al. 2006). The Framingham investigators subsequently have examined the association of the autosomal SNPs in relation to about 1000 phenotypes using generalized estimating equations (GEE) and family- based association tests (FBAT; Lange et al. 2003). The generalized estimating equation approach is a population-based strategy measuring association in a regression model that accounts for correlation among

related individuals. The FBAT procedure, on the hand, tests for differences in the probability of transmission of an allele based on phenotype from an expected Mendelian model and uses subsets of pedigrees that are informative for a SNP. Reflecting the complexity of the Framingham database, Framingham investigators have formed 17 phenotype-specific writing groups to examine these associations and publish the results. Plans are underway to replicate some of the findings either using “in silico” approaches or performing targeted association studies on other cohorts. Additionally, the Framingham investigators are collaborating with the National Center for Bioinformatics to develop a web display of the unfiltered results to speed data sharing and the ability to replicate our findings [**d**ata**b**ase of **G**enotype **a**nd **P**henotype, dbGaP; http://www.ncbi.nlm.nih.gov/SNP/GaP.html].

Genome-wide association studies present many challenges. The Framingham 100K genome-wide association studies have provided a window of opportunity to examine the complexities in organization and statistical analysis of these large data sets. Merely uploading, analyzing and synthesizing 100,000s of associations requires extensive resources and time. Interpreting the results has presented challenges. For example, should one use a minimum statistical significance (p-value) between a SNP and a known phenotype? Or should one use a complex phenotype or its components to perform association studies? In some instances different analytical approaches [genetic linkage, generalized estimating equations and family-based association testing] highlighted different SNPs and regions of interests. Distinguishing between true versus false positives in the context of 100,000s SNPs and hundreds of phenotypes has been daunting. The Framingham investigators have noted that most results are likely to be false positives and conversely, they may have failed to appreciate important true positives of modest statistical significance. Furthermore, these data provide additional raw material to understand the role of gene-gene interaction (both pleiotropy and epistatic gene action) and gene-environment interactions in the human genome and health. From this perspective, use of novel computational methods such as net work analysis and other machine learning approaches are contemplated.

The technology for genome-wide association studies has advanced rapidly, posing new ethical as well as analytical challenges. Framingham investigators work closely with three panels that deal with the ethical dimensions of genome-wide association studies: a) the Study’s Observational Safety and Monitoring Board, b) the Boston University Medical Center Institutional Review Board, and c) the Framingham Ethics Advisory Board. For instance, all the three panels have reviewed measures to protect participant confidentiality and ensure against genetic discrimination(Greely 2005; Billings 2005; Morrison 2005). In addition, the three panels are addressing under what circumstances it is appropriate to notify participants of the results of genetic testing (Bookman et al. 2006).

b. The SHARe (SNP-Health Association Resource) study: The National Heart Lung and Blood Institute has embarked on an ambitious collaboration with Boston University and Affymetrix to conduct a 550K genome-wide association study of 10,000 Original, Offspring and Third Generation Cohort participants and to post the aggregate results at the NCBI “dbGaP” (http://www.nih.gov/news/pr/dec2006/nlm-12.htm) website. Investigators around the world will be able to access the genotype and phenotype data collected over almost 6 decades after securing approval from the NHLBI, the scientist’s own Institutional Review Board, and signing a data distribution agreement. The objective is to speed scientific discovery while protecting Framingham participant confidentiality.

The genome-wide association studies at Framingham represent unparalleled opportunities as well as challenges. The challenges include bioinformatics, logistical, and ethical concerns. However, the extensive genotypic and phenotypic characterizations of Framingham

participants represent unique steps in the goal of achieving medical care that is “predictive, preemptive and personalized (Nabel 2006).

V. Genotype × environment interactions

The FHS has firmly established the role of environmental factors, such as the use of tobacco (Doyle *et al.*, 1992) and other life style factors (Posner *et al.*, 1993) on cardiovascular phenotypes. Since genes are known to interact with various environmental factors, their interaction may be reflected in the magnitude or in the direction of association. A number of polymorphisms in the candidate genes have been evaluated to determine their interaction with environmental factors. Some examples include: effects of dietary fatty acids on apolipoprotien A5 polymorphisms (Lai *et al.*, 2006); fatty acid binding protein (*FABP2*) in relation to plasma lipids (Galluzzi *et al.*, 2001); apolipoprotein E polymorphisms and alcohol consumption (Corella *et al.*, 2001). In an interesting study, Ordovas *et al.*, (2002) evaluated the relations between dietary fat intake and three genotypes of the C/T polymorphisms of the hepatic lipase gene (*LIPC*). They found a dose dependent association of T allele with higher HDL-C in subjects consuming <30 percent of the energy from fat (Figure 8). Also, the slopes formed by the genotypes in relation to gradient energy intake, followed the classical genotype × environmental interactions (Lynch and Walsh, 1998). These studies are providing valuable insights toward designing other large studies (Manolio *et al.*, 2006).

Prospects and conclusions

The Framingham Heart Study has made extraordinary contributions toward the discovery of cardiovascular risk factors and in turn has helped alleviate cardiovascular burden both in the US and elsewhere in the world. The availability of family structure and a rich panel of phenotypic data related to cardiovascular health as well as other ancillary traits are providing many useful insights on the role of genetic variation in cardiovascular risk traits, and their interaction with the environment. Interestingly, moderate to high heritability is common to many of the traits studied, suggesting a reservoir of genetic variation for the CV risk factors and other phenotypic traits. Also, heritability estimates vary over time or age among sexes. The longitudinal design and intensive phenotyping of the FHS participants increases the insights that may be obtained from the sample. For example, in this cohort, age can be matched and genetic variation can be measured over time to account for longitudinal changes in environmental factors affecting the trait of interest. Similarly, testing for the consistency of linkage peaks in relation to age or understanding pleiotropic gene action on seemingly different traits is facilitated by examining a sample such as the Framingham Heart study population. Answers obtained on genotype-environment interactions, using the FHS are already providing valuable insights toward designing additional studies and could further illuminate developing personalizing medications or interventions. Also, genome-wide association studies (e.g. Affymetrix116K chip) has made it possible to examine the genetic basis of numerous correlated traits and understand the challenges associated with such a large scale association study as well as examining the role of pleiotropy in the genome. The study is poised to perform association studies using the Affymetrix 550k chip. This effort should provide additional insights toward refining the locations of candidate and novel genes, as well as to ask other questions relating to functional aspects of the identified genes. Answers to these fundamental questions may hold promise toward applying genetics and evolutionary principles to both public health and to the practice of medicine.

Acknowledgments

The investigators are deeply appreciative to the three generations of Framingham Heart Study participants. DRG is thankful to Ms. Lynnel Lyons and Ms. Esta Shindler for their help in the literature survey. We also thank the core contract **NO1-HC 25195, and RO1s HL076784, AG 028321**.

References

Altshuler D, Brtooks LD, Chakravarthy A, Collins FS, Daly MJ, Donnelly P. A haplotype map of the human genome. Nature 2005;437:1299–1320. [PubMed: 16255080]

Anonymous. Familial aggregation of lens opacities: The Framingham eye study and the Framingham offspring study. Am. J. Epid 1994;140:555–564.

Anonymous. A century of medical milestones: Editors of Merck Manual assess the most important advances of the last 100 years. Washington Post: 1999 Dec 21.

Ardlie KG, Kruglyak L, Seilstand M. Pattern of linkage disequilibrium in the human genome. Nat. Rev. Genet 2002;3:299–309. [PubMed: 11967554]

Arnesen E, Refsum H, Bonaa KH, Ueland PM, Forde OH, Nordrehaug JE. Serum total homocysteine and coronary heart disease. Int J Epidemiol 1995;24:704–709. [PubMed: 8550266]

Arya R, Lehman D, Hunt KJ, Schneider J, Almasy L, Blangero J, Stern MP, Duggiral R. Evidence for bivariate linkage of obesity and HDL-C levels in the Framingham Heart Study. BMC Genetics 2003;4:S52. [PubMed: 14975120]

Atwood LD, Heard-Costa NL, Cupples LA, Jaquish CE, Wilson PWF, D'Agostino RB. Genomewide linkage analysis of body mass index across 28 years of the Framingham Heart Study. Am. J. Hum. Genet 2002;71:1044–1050. [PubMed: 12355400]

Atwood L, Wolf PA, Heard Costa NL, Massaro JM, Beiser A, D'Agostino RB, DeCarli C. Genetic variation in white matter hyperintensity volume in the Framingham Study. Stroke 2005;35:1609–1613. [PubMed: 15143299]

Atwood LD, Heard-Costa NL, Fox CS, Jaquish CE, Cupples LA. Sex and age specific effects of chromosomal regions linked to body mass index in the Framingham Study. BMC Genet 2006;7:7–?. [PubMed: 16438729]

Badyayev AV, Foresman KR. Evolution of morphological integration: I. Functional units channel stress-induced variation. Am. Natur 2004;163:868–879. [PubMed: 15266384]

Benajmin EJ, Larson MG, Keys MJ, Mitchell GF, Vasan RS, Keanney JF, Lehman BT, Fan S, Osypiuk E, Vita JA. Clinical correlates and heritability of flow-mediated dilation in the community: Te Framingham Heart Study. Circulation 2004;190:613–619.

Billings PR. Genetic nondiscrimination. Nat. Genet 2005;37:559–560. [PubMed: 15920511]

Bookman EB, Langehorne AA, Eckfeldt JH, Glass KC, Jarvik GP, Klag M, Koski G, Motulsky A, Wilfond B, Manolio TA, Fabsitz RR, Luepker RV. NHLBI working Group Reporting genetic results in research studies: summary and recommendations of an NHLBI working group. Am J Med. Genet 140:1033–1040.

Botstein D, White RL, Skolnik M, Davis MW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet 1980;32:314–321. [PubMed: 6247908]

Brand FN, Kiely DK, Kannel WB, Myers RH. Family patterns of coronary heart disease mortality: The Framingham longevity study. J. Clin. Epidemiol 1992;45:169–174. [PubMed: 1573433]

Brousseau M, Bodzioch M, Schafer EJ, Goldkamp AL, Kielar D, Probst M, Ordovas JM, Aslanidis C, Lackner KJ, Rubin HB, Collins D, Robins SJ, Wilson PWF, Scmitz G. Common variants in the gene encoding ATP-binding cassette transporter 1 in men with low HDL cholesterol levels and coronary heart disease. Atherosclerosis 2001;154:607–611. [PubMed: 11257261]

Brown WM, Beck SR, Lange EM, Davis CC, Kay CM, Langefield CD, Ricj SS. Age stratified heritability estimation in the Framingham Heart Study families. BMC Genetics 2003;4:S32. [PubMed: 14975100]

Butler RN. Framingham: The town with heart (editorial). Geriatrics 1991;54:3–4.

Carlson CS, Eberle MA, Kruglyak L, Nickerson DA. Mapping complex disease loci in whole-genome association studies. Nature 2004;429(6990):446–452. [PubMed: 15164069]

Cheverud JM. A comparison of genetic and phenotypic correlations. Evolution 1988;42:958–968.

Cheverud, JM. The genetic architecture of pleiotropic relations and differential epistasis. In: Wagner, GP., editor. The Character Complex in Evolutionary Biology. New York, NY: Academic Press; 2001.

Churchill, GA. Co-adapted alleles and evolution of complex traits; Abstract Presented at the Keystone Symposium; Big Sky, MT: 2006 Jan 7–12.

Cooper R, Cutler J, Desvigne-Nickens P, Fortman SP, Friedman L, Havlik R, Hogelin G, Marler J, McGovern P, Morosco G, Mosca L, Pearson T, Stamler J, Stryer D, Thom T. Trends and disparities in coronary heart disease, stroke, and other cardiovascular diseases in the United States. Circulation 2000;102:3137–3147. [PubMed: 11120707]

Corella D, Tucker K, Lahoz C, Coltell O, Cupples LA, Wilson PW, Schaefer EJ, Ordovas JM. Alcohol drinking determines the effect of the APOE locus on LDL-cholesterol concentrations in men: the Framingham Offspring Study. Am J Clin Nutr 2001;73:736–745. [PubMed: 11273848]

Dawber, TR.; Kannel, WB. Studies of historical interest. In: Levy, D., editor. 50 Years of discovery: medical milestones from the National Heart, Lung, and Blood Institute's Framingham Heart Study. Hackensack, NJ: Center for Biomedical Communications, Inc.; 1999. p. 3-4.

Dawber TR, Kannel WB, Revotskie N, Stokes J III, Kagan A, Gordon T. Some factors associated with the development of coronary heart disease. Amer. J. Public Health 1959;49:1349–1356.

Dawber TR, Meadors GF, Moore FE Jr. Epidemiological approaches to heart disease: The Framingham Study. Am. J. Public Health 1951;41:279–286.

Demisse S, Cupples LA, Myers R, Aliabadi P, Levy D, Felson DT. Genome scan for quantity of hand osteoarthritis. Arthritis Rheum 2002;46:946–952. [PubMed: 11953971]

Devlin B, Risch N. A comparison of linkage disequilibrium measures for fine scale mapping. Genomics 1995;29:311–322. [PubMed: 8666377]

Doyle JT, Dawber TR, Kannel WB, Heslin AS, Kahn HA. Cigarette smoking and coronary heart disease. N. Engl. J. Med 1962;266:796–801. [PubMed: 13887664]

Dupius J, Larson MG, Vasan RS, Massaro JM, Wilson PW, Lipinska I, Corey D, Vita JA, Keaney JF, Benjamin EJ. Genome scan of systemic biomarkers of vascular inflammation in the Framingham Heart Study: evidence for susceptibility loci on 1q. Atherosclerosis 2005;182:307–314. [PubMed: 16159603]

Elias MF, Sullivan LM, D'Agostino RB, Elias PK, Jacques PF, Selhub J, Seshadri S, Au R, Beiser A, Wolf PA. Homocysteine and cognitive performance in the Framingham offspring study: age is important. 2005;162:644–653.

Fabsitz RR, Nam J-M, Gart J, Sttunkard A, Price AR, Wilson PWF. HLA associations with obesity. Hum. Hered 1989;39:156–164. [PubMed: 2591979]

Falconer, DS.; Mackay, FC, Tr. Introduction to Quantitative Genetics. Essex, U. K: Longman; 1996.

Felson DT, Couropmitree NN, Chaisson CE, Hannan MT, Zhang Y, McAlindon TE, LaValley M, Levy D, Myers RH. Evidence for a Mendelian gene in a segregation analysis of generalized radiographic osteoarthritis: the Framingham Study. Arthritis Rheum 1998;41:1064–1071. [PubMed: 9627016]

Fox CS, Polak JF, Chazaro I, Cupples A, Wolf PA, D'Agostino RA, O'Donnell CJ. Genetic and environmental contributions to atherosclerosis phenotypes in men and women: heritability of carotid intima-media thickness in the Framingham Heart Study. Stroke 2003;34:397–401. [PubMed: 12574549]

Fox CS, Evans JC, Larson M, Kannel WB, Levy D. Temporal trends in coronary heart disease mortality and sudden cardiac death from 1951 to 1999. Circulation 2004;110:522–527. [PubMed: 15262842]

Fox CS, Cupples LA, Chazaro I, Polak JF, Wolf PA, D'Agostino RB, Ordovas JM, O'Donnell CJ. Genomewide linkage analysis for internal carotid artery intimal medial thickness: evidence for linkage to chromosome 12. Am J Hum Genet 2004;74:253–261. [PubMed: 14730480]

Fox CS, Heard-Costa NL, Wilson PWF, Levy D, D'Agostino RB Sr, Atwood LD. Genome-wide linkage to chromosome 6 for waist circumference in the Framingham Heart Study. Diabetes 2004;53:1399–1402. [PubMed: 15111512]

Fox CS, Heard-Costa NL, Vasan RS, Murabito JM, D'Agostino RB, Atwood LD. Genomewide linkage analysis of weight change in the Framingham Heart Study. J. Clin. Endocrinol. Metab 2005;90:3197–3201. [PubMed: 15769990]

Friedlander Y, Kark JD, Stein Y. Family history as an independent risk factor for coronary heart disease. Br. Heart J 1985;53:382. [PubMed: 3986055]

Galluzzi JR, Cupples LA, Meigs JB, Wilson PFW, Shafer EJ, Ordovas JM. Association of the ala54-thr polymorphism in the intestinal fatty acid-binding protein with 2-h postchallenge insulin levels in the Framingham Offspring Study. Diabetes Care 2001;24:1161–1166. [PubMed: 11423496]

Galton F. Family likeness in stature. Proc. Roy. Soc. Lond 1886;40:42–63.

Galton, F. English Men of Science: Their Nature and Nurture. New York: D. Appleton and Company; 1890.

Galton F. Why do we measure mankind? Lippincott's monthly magazine 1890;45:236–241.

Garrod AE. The incidence of Alkaptonuria: A study in chemical individuality. Lancet 1902;2:1616–1620.

Givelber RJ, Couropmitree NN, Gottlieb DJ, Evans JC, Levy D, Myers RH, O'Conner GT. Segregation analysis of pulmonary function among families in the Framingham Study. Am. J. Respir. Care Med 1998;157:1445–1451.

Gordon, T.; Kannel, WB. The Framingham, Massachusetts, Study Twenty years later. In: Kessler, II.; Levin, ML., editors. The Community as an Epidemiologic Laboratory; A Casebook of Community Studies. Baltimore, MD: Johns Hopkins University Press; 1970. p. 123-146.

Greely HT. Banning genetic discrimination. N.Engl.J Med 2005;353:865–867. [PubMed: 16135828]

Guo Z, Li X, Rao S, Moser KL, Zhang T, Gong B, Shen G, Li L, Cannata R, Zirzow E, Topol EJ, Wang Q. Multivariate sib-pair linkage analysis of longitudinal phenotypes by three step-wise analysis approaches. BMC Genet 2003;4(1):S68. [PubMed: 14975136]

Harris H. Enzyme polymorphisms in man. Proc. R. Soc. Lond. B. Biol. Sci 1966;164:298–310. [PubMed: 4379519]

Havlik RJ, Garrison RJ, Feinleib M, Kannel WB, Castelli WP, McNamara PM. Blood pressure aggregation in families. Am. J. Epidemiol 1979;110:304–312. [PubMed: 474567]

Herbert A, Gerry NP, McQueen MB, Heid IM, Pfeufer A, Illig T, Wichmann HE, Meitinger T, Hunter D, Hu FB, Colditz G, Hinney A, Hebebrand J, Koberwitz K, Zhu X, Cooper R, Ardlie K, Lyon H, Hirschhorn JN, Laird NM, Lenburg ME, Lange C, Christman MF. A common genetic variant is associated with adult and childhood obesity. Science 2006;312:279–283. [PubMed: 16614226]

Hill WG, Weir BS. Maximum likelihood estimation of gene location by linkage disequilibrium. Am. J. Hum. Genet 1994;54:705–714. [PubMed: 8128969]

Hubert HB, Feinleib, McNamara PM, Castelli WP. Obesity as an independent risk factor for cardiovascular disease: A 26-year follow up of participants in the Framingham Heart Study. Circulation 1983;67:968–977. [PubMed: 6219830]

Ioannidis JP, Trikalinos TA, Ntzani EE, Contopoulos-Ioannidis DG. Genetic associations in large versus small studies: an empirical assessment. Lancet 2003;361:567–571. [PubMed: 12598142]

Jeffreys AJ, Wilson V, Thein SL. Hypervariable 'minisatellite' regions in human DNA. Nature 1985;314:467–473. [PubMed: 2984578]

Kannel WB. Contribution of the Framingham Study to preventive cardiology. J. Am. Coll. Cardiol 1990;15:206–211. [PubMed: 2136875]

Kannel WB, Ellison RC. Alcohol and coronary heart disease: the evidence for a protective effect. Clinica Chimica Acta 1996;246:59–76.

Kannel WB, Dawber TR, Kagan A, Revotskie N, Stokes J III. Factors of risk in the development of coronary heart disease – six-year follow-up experience. Ann. Int. Med 1961;55:33–50. [PubMed: 13751193]

Kagan A, Dawber TR, Kannel WB, Revotskie N. The Framingham Study: a perspective study of coronary heart disease. Federation Proceedings 1962;21:52–57. [PubMed: 14453051]

Kannel WB, Gordon T, Offut D. Left ventricular hypertrophy by electrocardiogram: Prevalence, incidence, and mortality in the Framingham Study. Ann. Int. Med 1969;71:89–105. [PubMed: 4239887]

Kannel WB, Feinleib M, McNamara PA, Garrison RJ, Castelli WP. An investigation of coronary hear disease in families: the Framingham offspring study. Amer J. Epid 1979;110:282–290.

Kannel, WB.; Stokes, JI. The epidemiology of coronary artery disease. In: Cohn, PF., editor. Diagnosis and Therapy of Coronary Artery Disease. Boston, MA: Martinus Nijhoff; 1985. p. 63-88.

Kannel WB, Wolf PA, Castelli WP, D'Agostino RB. Fibrinogen and risk of cardiovascular disease: The Framingham Study. JAMA 1987;258:1183–1186. [PubMed: 3626001]

Karasik D, Myers RH, Hannan MT, Gagnon D, McLean RR, Cupples LA, Kiel D. Genome screen for quantitative trait loci contributing to normal variation in bone mineral density: The Framingham Study. J. Bone Miner. Res 2002;17:1718–1727. [PubMed: 12211443]

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;33:308–316. [PubMed: 13678771]

Kathiresan S, Larson MG, Vasan RS, Guo C-Y, Gona P, Keaney JF, Wilson PWF, Newton-Cheh C, Musone SL, Camargo AL, Drake JA, Levy D, O'Donnell CJ, Hirschhorn JN, Benjamin EJ. Contribution of clinical correlates and 13 C-reactive protein gene polymorphisms to interindividual variability in serum C-Reactive protein level. Circulation 2006;113:1415–1423. [PubMed: 16534007]

Keaney JF, Massaro JM, Larson MG, Vasan RS, Wilson PWF, Lipinska I, Corey D, Sutherland P, Vita JA, Benjamin EJ. Heritability and correlates of intercellular adhesion molecule-1 in the Framingham offspring study. J. Am. Coll. Cardiol 2004;44:168–173. [PubMed: 15234428]

Klein RJ, Zeiss C, Chew EY, Tsai JY, Sackler RS, Haynes C, Henning AK, SanGiovanni JP, Mane SM, Mayne ST, Bracken MB, Ferris FL, Ott J, Barnstable C, Hoh J. Complement factor H polymorphism in age-related macular degeneration. Science 2005;308:385–389. [PubMed: 15761122]

Lai C_Q, Corella D, Demisse S, Cupples LA, Adiconis X, Zhu Y, Parnell L, Tucker KL, Ordovas JM. Dietary intake of n-6 fatty acids modulates effect of apolipoprotien A5 gene on plasma fasting triglycerides, remnant lipoprotein concentrations, and lipoprotien particle size. The Framingham Study. Circulation 2006;113:2062–2070. [PubMed: 16636175]

Lange C, Silverman EK, Xu X, Weiss ST, Laird NM. A multivariate family-based association test using generalized estimating equations: BAT-GEE. Biostatistics 2003;4:196–206.

Lee DS, Pencina MJ, Benjamin EJ, Wang TJ, Levy D, O'Donnell CJ, Nam B-H, Larson MG, D'Agostino RB, Vasan RS. Association of parental heart failure with risk of heart failure in offspring. N. Engl. J. Med 2006;355:138–147. [PubMed: 16837677]

Levy, D.; Brink, S. A Change of Heart: How the people of Framingham, Massachusetts, helped unravel the mysteries of cardiovascular disease. New York, NY: Alfred A. Knopf; 2005.

Levy D, DeStefano AL, Larson MG, O'Donnell CJ, Lifton RP, Gavras H, Cupples LA, Myers RH. Evidence for a gene influencing blood pressure on chromosome 17: genome scan linkage results for longitudinal blood pressure phenotypes in subjects from the Framingham Hear Study. Hypertension 2000;36:477–483. [PubMed: 11040222]

Levy D, DePalma SR, Benjamin E, O'Donnell CJ, Parise H, Hirschhorn JN, Vasan RS, Izumo S, Larson MG. Phenotype-genotype association grid: a convenient method for summarizing multiple association analyses. BMC Genetics 2006;7:30. [PubMed: 16716207]

Lewontin RC. The interaction of selection and linkage. I. General consideration, heterotic models. Genetics 1964;49:49–67. [PubMed: 17248194]

Lin JP, O'Donnell CJ, Levy D, Cupples LA. Evidence for a gene influencing haematocrit on chromosome 6q23-24: genomewide scan in the Framingham Heart Study. J. Med. Genet 2005;42:75–79. [PubMed: 15635079]

Lin J-P. Genome-wide scan on plasma triglycerides and high density lipoproteins cholesterol levels, accounting for the effects of correlated quantitative phenotypes. BMC Genetics 2003;4:S47. [PubMed: 14975115]

Liu X-Q, Hanley AJG, Paterson AD. Genetic analysis of common factors underlying cardiovascular disease related traits. BMC Genetics 2003;4:S56. [PubMed: 14975124]

Lohmueller KE, Pearce CL, Pike M, Lander ES, Hirschhorn JN. Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. Nat.Genet 2003;33:177–182. [PubMed: 12524541]

Lloyd-Jones DM, Nam B-H, D'Agostino RB, Levy D, Murabito JM, Wang TJ, Wilson PWF, O'Donnell CJ. Parental cardiovascular disease as a risk factor for cardiovascular disease in middle-age adults. JAMA 2004;291:2204–2211. [PubMed: 15138242]

Lynch, M.; Walsh, B. Genetics and Analysis of Quantitative Traits. Sunderland, MA: Sinauer Associates; 1998.

Mackay TFC, Lyman RF. Drosophila bristles and the nature of quantitative genetic variation. Phil. Trans. R. Soc 2005;360:1513–1527.

Manolio TA, Bailey-Wilson JE, Collins FS. Genes, environment and the value of prospective cohort studies. Nature Rev. Genet 2006;7:812–820. [PubMed: 16983377]

Martin LJ, North KE, Dyer T, Blangero J, Comuzzie AG, Williams J. Phenotypic, genetic, and genome-wide structure in the metabolic syndrome. BMC Genet 2003;4:S95. [PubMed: 14975163]

Mehta N, Khan IA. Cardiology's 10 greatest discoveries of the 20th century. Tx. Heart Inst. J 2002;29:164–171.

Mitchell GF, DeStefano AL, Larson MG, Benjamin EJ, Chen MH, Vasan RS, Vita JA, Levy D. Heritability and a genome-wide linkage scan for arterial stiffness, wave reflection, and mean arterial pressure: the Framingham Heart Study. Circulation 2005;112:194–199. [PubMed: 15998672]

Morrison PJ. Insurance, unfair discrimination, and genetic testing. Lancet 2005;366:877–880. [PubMed: 16154000]

Murabito JM, Yang Q, Fox CS, Cupples LA. Genome-wide linkage analysis to age at natural menopause in a community-based sample: the Framingham Heart Study. Fertil. Steril 2005;84:1674–1679. [PubMed: 16359963]

Myers RH, Kiley DK, Cupples LA, Kannel WB. Parental history is an independent factor for coronary heart disease: The Framingham Heart Study. Am. Heart J 1990;120:963–969. [PubMed: 2220549]

Nabel, EG. Genomic medicine and cardiovascular disease. Simon Dack Lecture. The American College of Cardiology. 2006. http://www.nhlbi.nih.gov/directorspage/pageimages/03-11-06-acc_dack_nabel.pdf

Newton-Cheh C, Larson MG, Corey DC, Benjamin EJ, Herbert AG, Levy D, D'Agostino RB, O'Donnell CJ. QT interval is a heritable quantitative trait with evidence of linkage to chromosome 3 in a genome-wide linkage analysis: The Framingham Heart Study. Heart Rhythm 2004;2:277–284. [PubMed: 15851319]

O'Donnell CJ, Larson MG, Feng D, Sutherland PA, Lindpaintner K, Myers RH, D'Agostino RA, Levy D, Tofler GH. Genetic and environmental contributions to platelet aggregation: the Framingham heart study. Circulation 2001;103:3051–3056. [PubMed: 11425767]

O'Donnell CJ, Chazaro I, Wilson PW, Fox C, Hannan MT, Kiel DP, Cupples LA. Evidence for heritability of abdominal aortic calcific deposits in the Framingham Heart Study. Circulation 2002;106:337–341. [PubMed: 12119250]

Ordovas JM, Civeira F, Genest J Jr, Craig S, Robbins AH, Meade T, Pocovi M, Frosaard PM, Masharani U, Wilson PWF, Salem DN, Ward RH, Shaefer EJ. Restriction fragment length polymorphisms of the apolipoprotien A-I, C-III, A-IV gene locus. Atherosclerosis 1991;87:75–86. [PubMed: 1678604]

Ordovas JM, Corella D, Demissie S, Cupples LA, Couture P, Coltell O, Wilson PWF, Scahefer EJ, Tucker KL. Dietary fat intake determines the effect of common polymorphism in the Hepatic Lipase gene promoter on high-density lipoprotein metabolism. Evidence of a strong dose effect in this gene-nutrient interaction in the Framingham Study. Circulation 2002;106:2315–2321. [PubMed: 12403660]

Peter I, Shearman A, Vasan RS, Zucker DR, Schmid CH, Demissie S, Cupples AL, Kuvin JT, Karas RH, Mendelsohn ME, Housman DE, Benjamin EJ. Association of estrogen receptor β gene

polymorphisms with left ventricular mass and wall thickness in women. Am. J. Hepertens 2005;18:1388–1395.

Posner BM, Cupples LA, Gagnon D, Wilson PWF, Chetwynd K, Felix D. Healthy people 2000. The rationale and potential efficacy of preventive nutrition in heart disease: The Framingham Offspring-Spouse Study. Arch. Intern. Med 1993;153:1549–1556. [PubMed: 8391793]

Post WS, Larson MG, Myers RH, Galderisi M, Levy D. Heritability of left ventricular mass: The Framingham Heart Study. Hypertension 1997;30:1025–1028. [PubMed: 9369250]

Reed T, Kirkwood SC, DeCarli C, Swan GE, Miller BC, Wolf PA, Jack LM, Carmelli D. Relationship of family history scores for stroke and hypertension to quantitative measures of white-matter hyperintensities and stroke volume in elderly males. Neuroepidemiology 2000;19:76–86. [PubMed: 10686532]

Rice JP, Saccone NL, Rasmussen E. Definition of the Phenotype. Adv. Genet 2001;42:69–78. [PubMed: 11037314]

Risch N, Merikangas K. The future of genetic studies of complex human diseases. Science 1996;273:1516–1517. [PubMed: 8801636]

Sham, P. Statistics in Human Genetics. New York, NY: Arnold; 1988.

Seman LJ, DeLuca C, Jenner JL, Cupples LA, McNamara JR, Wilson PWF, Castelli WP, Ordovas JM, Schaeffer EJ. Lipoprotien (a)- cholesterol and coronary heart disease in the Framingham Heart Study. Clinical Chem 1999;45:1039–1046. [PubMed: 10388480]

Shearman AM, Cupples LA, Demissie S, Peter I, Schmid CH, Karas CH, Mendelsohn ME, Housman DE, Levy D. Association between Estrogen receptor a gene variation and cardiovascular disease. JAMA 2003;290:2263–2270. [PubMed: 14600184]

Sheshadri S, Beiser A, Selhub J, Jacques PF, Rosenberg IH, D'Agostino RB, Wilson PWF, Wolf PA. Plasma homocysteine as a risk factor for dementia and Alzheimer's disease. N. Engl. J. Med 2002;346:476–483. [PubMed: 11844848]

Singh JP, Larson MG, O'Donnell CJ, Tsuji H, Evans JC, Levy D. Heritability of heart rate variability: The Framingham Heart Study. Circulation 1999;99:2251–2254. [PubMed: 10226089]

Solomon E, Bodmer WF. Evolution of sickle cell variant gene. Lancet 1979;333:923. [PubMed: 86686]

Strachan, T.; Read, AP. Human Molecular Genetics. New York, NY: Wiley-Liss; 2003.

Tai ES, Demisse S, Cupples LA, Corella D, Wilson PF, Schaefer EJ, Ordovas JM. Association between the PPARA L162V polymorphism and plasma lipid levels. The Framingham Offspring study. Arterioscler Thromb Vasc. Biol 2002;22:805–810. [PubMed: 12006394]

Terwilliger, JD.; Ott, J. Handbook of Human Genetic Linkage. Baltimore, MD: Johns Hopkins University Press; 1994.

Thom T, Haase N, Rosamond W, Howard VJ, Rumsfeld J, Manolio T, Zheng ZJ, Flegal K, O'Donnell C, Kittner S, Lloyd-Jones D, Hong Y. Heart disease and stroke statistics--2006 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Circulation 2006;113:e85–e151. [PubMed: 16407573]

Truett J, Cornfield J, Kannel W. A multivariate analysis of the risk of coronary heart disease in Framingham. J. Chron. Dis 1967;20:511–524. [PubMed: 6028270]

Vasan RS. Biomarkers of cardiovascular disease: molecular basis and practical considerations. Circulation 2006;113:2335–2362. [PubMed: 16702488]

Wang TJ, Nam B-H, Wilson PWF, Wolf PA, Levy D, Polak JF, D'Agostino RB, O'Donnell CJ. Association of C-reactive protein with carotid atherosclerosis in men and women: The Framingham Heart Study. Arterioscler Thromb Vasc Biol 2002;22:1662–1667. [PubMed: 12377746]

Wang TJ, Larson MG, Levy D, Benjamin EJ, Corey D, Leip EP, Vasan RS. Heritability and genetic linkage of plasma natriuretic peptide levels. Circulation 108:13–16. [PubMed: 12821537]

Weber JL, May PE. Abundant class of human DNA polymorphisms which can be typed using polymerase chain reaction. A,. J. Hum. Genet 1989;44:388–396.

Wilk JB, DeStefano AL, Arnett DK, Rich SS, Djousse L, Crapo RO, Leppert MF, Province MA, Cupples LA, Gottlieb DJ, Myers RH. A genome-wide scan of pulmonary function measures in

the National Heart, Lung, and Blood Institute Family Heart Study. Am J Respir Crit Care Med 2003;167:1528–1533. [PubMed: 12637344]

Wilson PWF. Established risk factors and coronary heart disease: The Framingham Study. AJH 1994;7:7S–12S. [PubMed: 7946184]

Wilson PWF, Myers RH, Larson MG, Ordovas JM, Wolf PA, Schafer EJ. Apolipoprotien e alleles, dyslipidemia, and coronary heart disease. The Framingham Offspring Study. JAMA 1994;272:1666–1671. [PubMed: 7966894]

Wilson PWF, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. Circulation 1998;97:1837–1847. [PubMed: 9603539]

Yang Q, Lai CQ, Parnell L, Cupples LA, Adiconis X, Zhu Y, Wilson PW, Housman DE, Shearman AM, D'Agostino RB, Ordovas JM. Genome-wide linkage analyses and candidate gene fine mapping for HDL3 cholesterol: the Framingham Study. J. Lipid. Res 2005;46:1416–1425. [PubMed: 15805549]

Yusuf S, Reddy S, Ounpuu S, Anand S. Global burden of cardiovascular diseases. Part II. Variations in cardiovascular disease by specific ethnic groups and geographic regions and prevention strategies. Circulation 2001;104:2855–2864. [PubMed: 11733407]

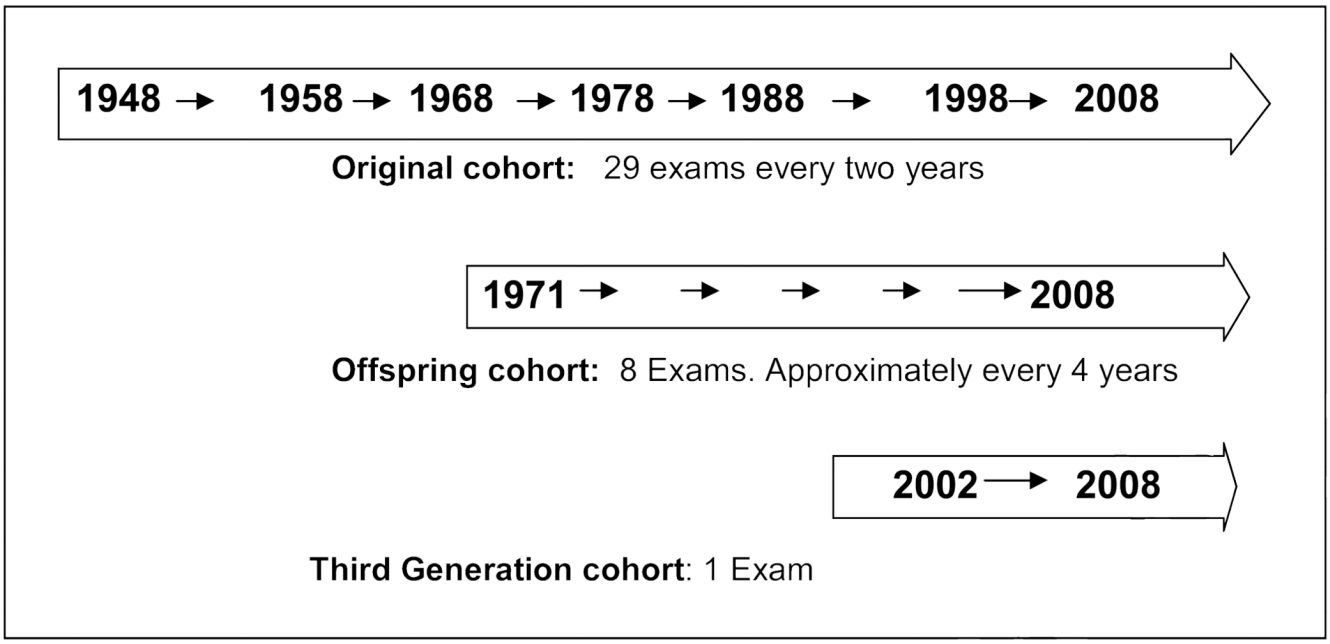


Figure 1.
Initiation and progression of examinations among three generations of participants in the Framingham Heart Study

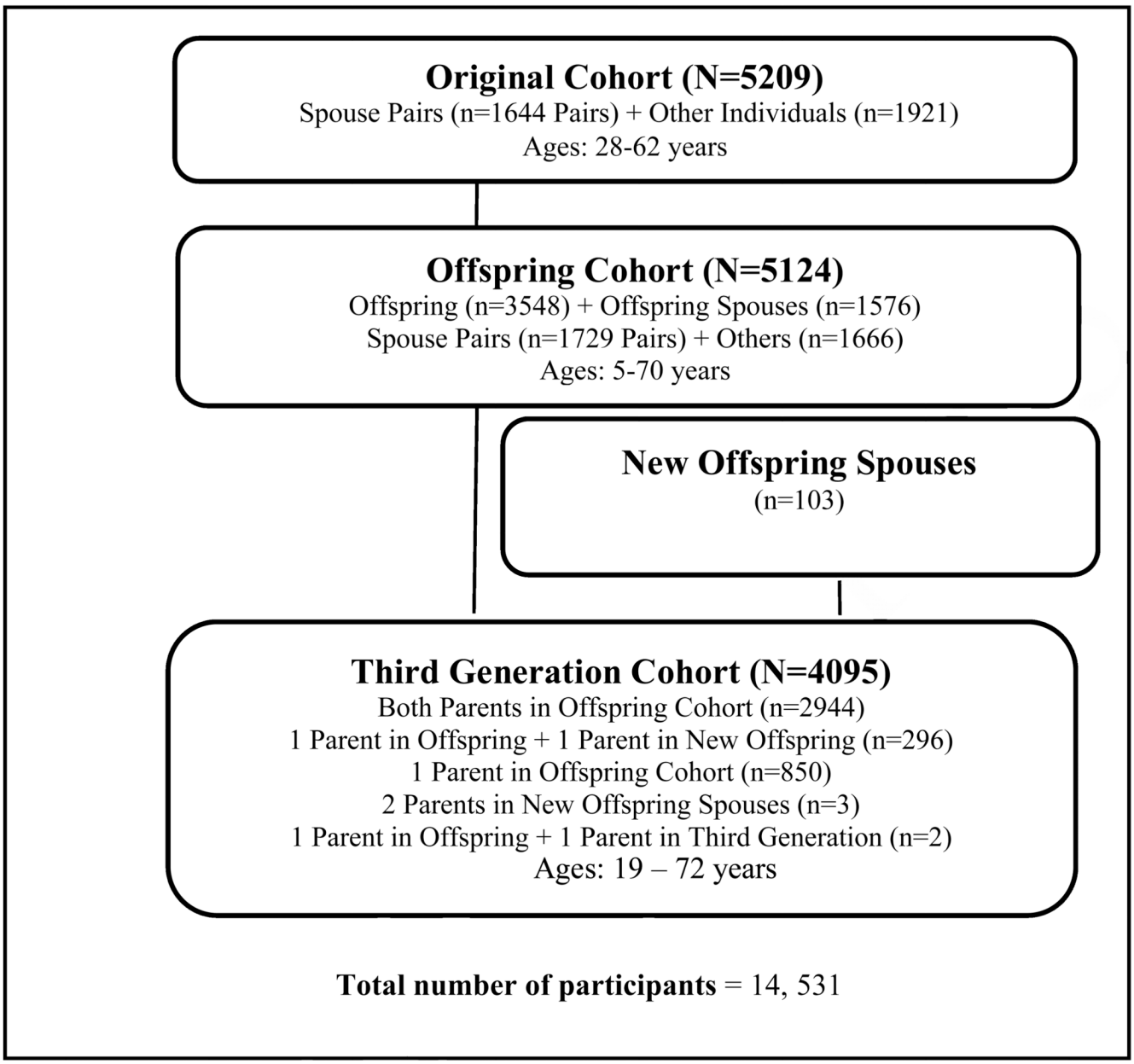


Figure 2.
Distribution of participants in each of the three generations in the Framingham Heart Study

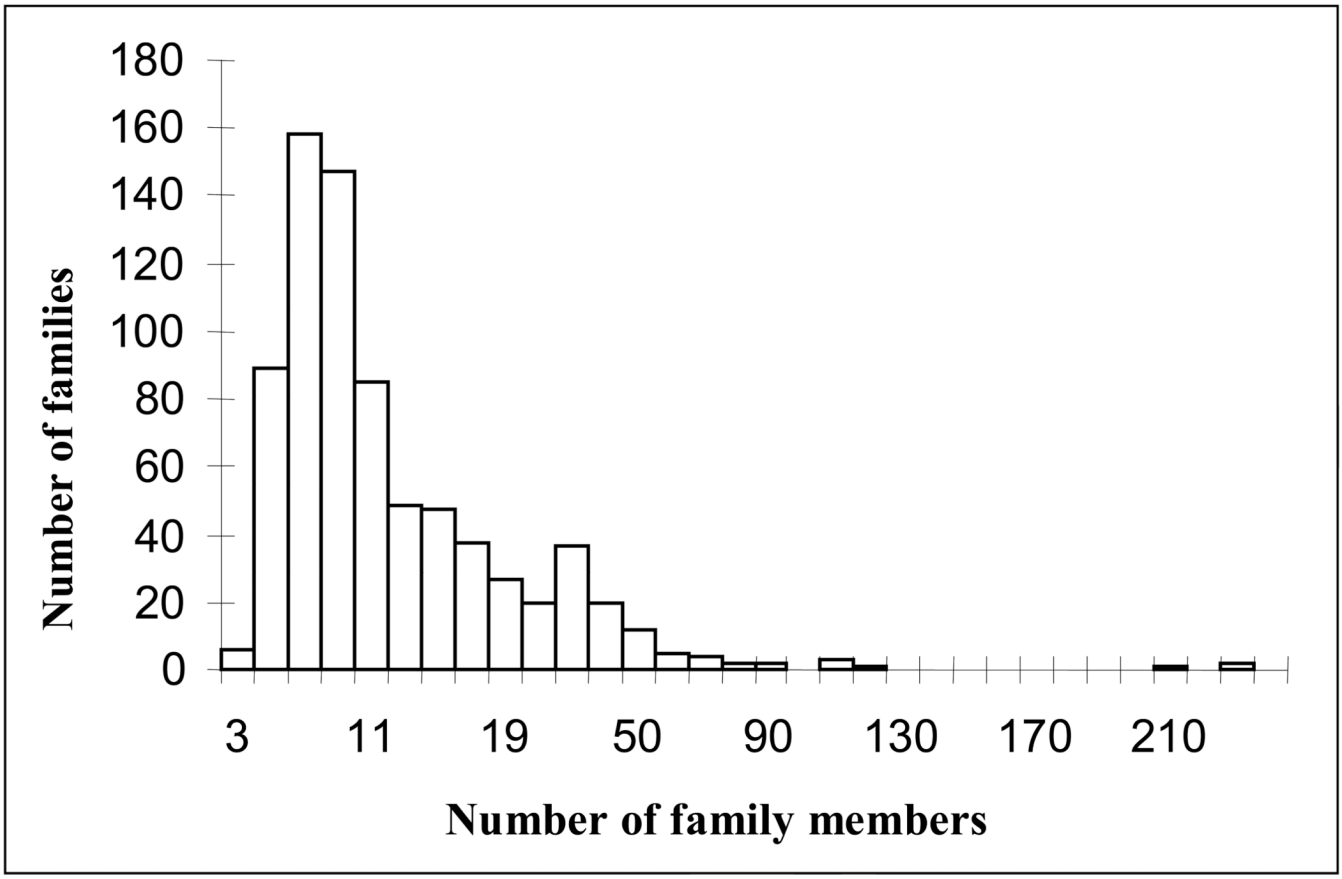


Figure 3.
Distribution of families in relation to pedigree sizes in the Framingham Heart Study

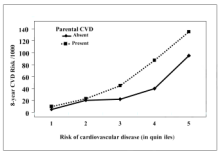


Figure 4.
Cardiovascular risk between parents and offspring in relation to quintiles of major risk factors: systolic blood pressure, body mass index, total to high density lipoprotein-cholesterol, diabetes and smoking (Lloyd-Jones et al. 2004).

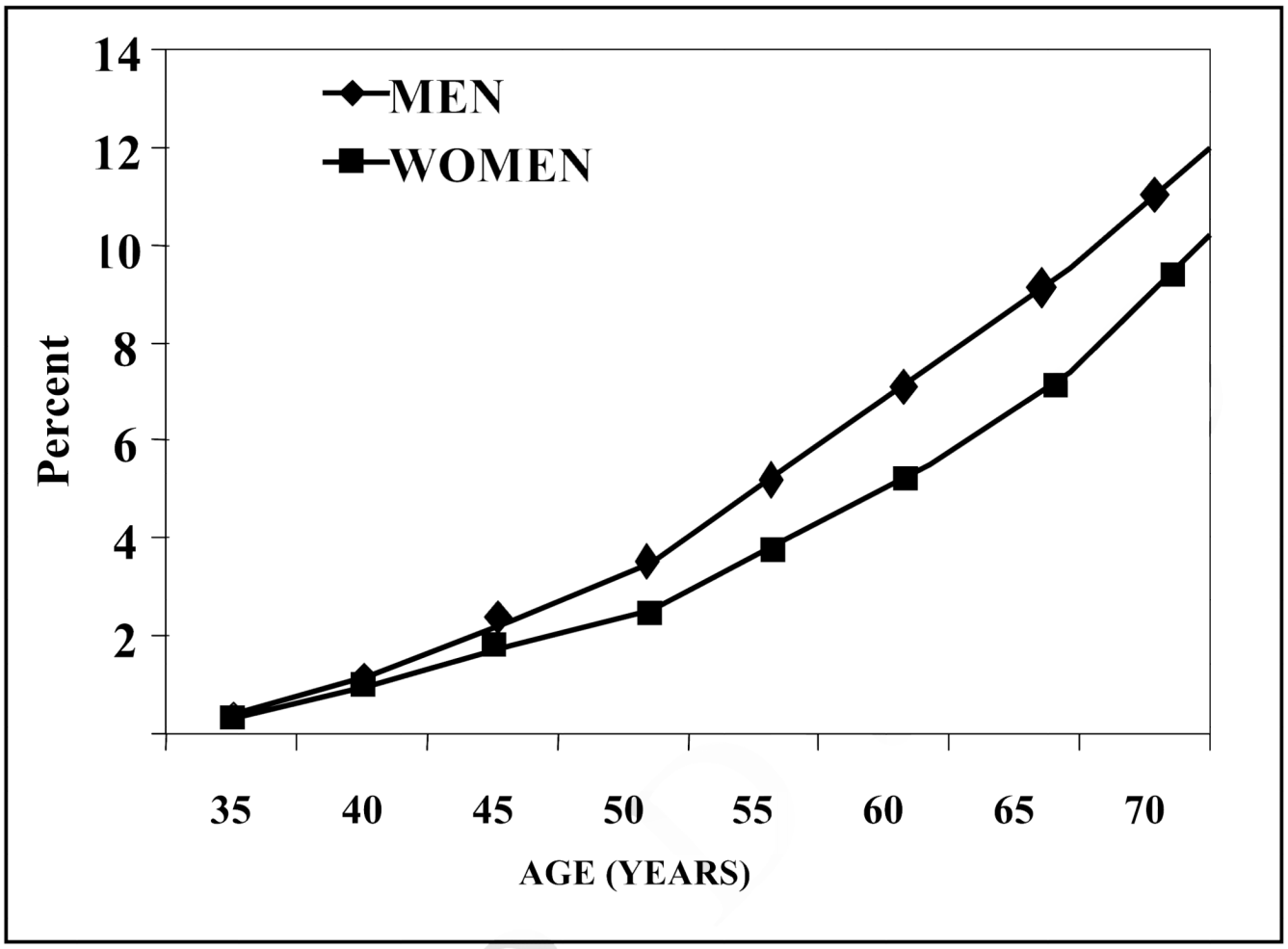


Figure 5.
Sex difference in susceptibility to cardiovascular diseases over 26 years in the Framingham Heart Study population (Hubert et al 1983).

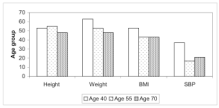


Figure 6.
Variation of heritability across age groups among four traits (Brown *et al*., 2003)

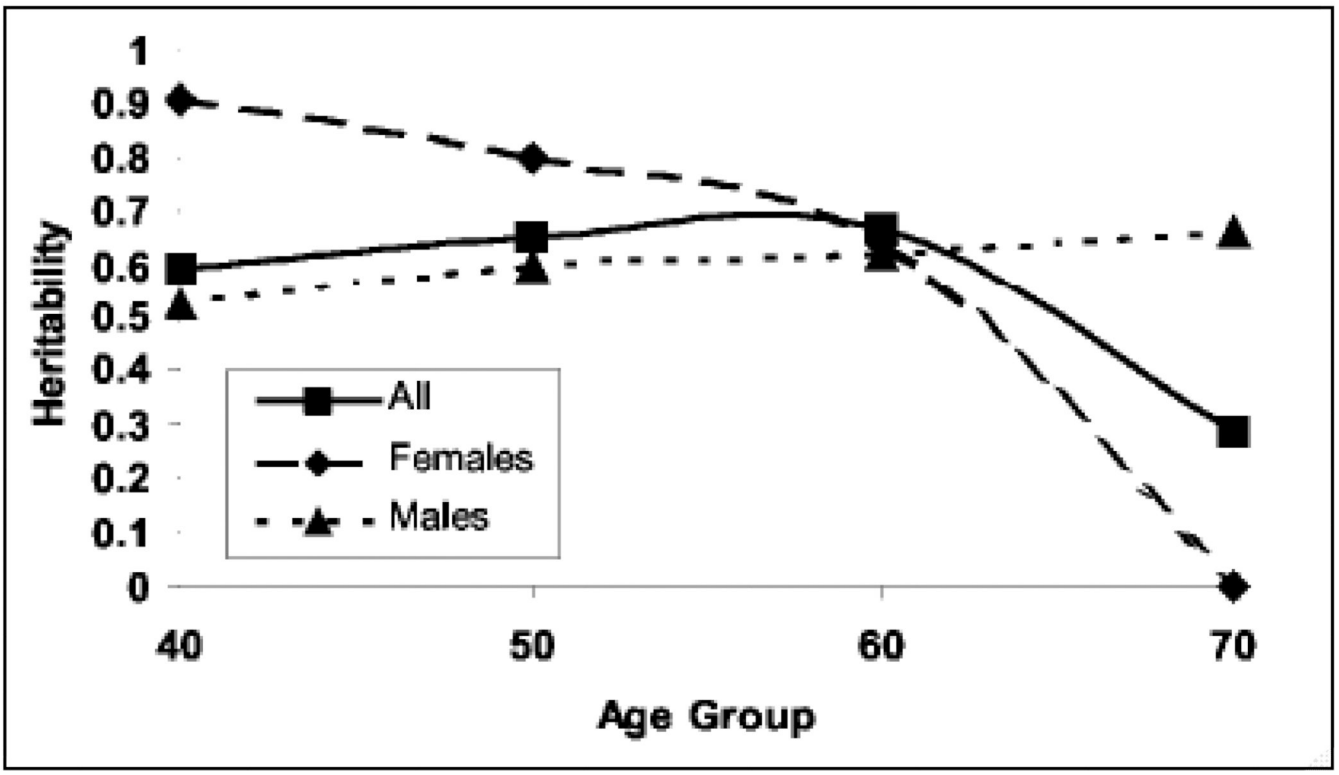


Figure 7.
Variation of heritability for white matter hypersensitivity volume between males and females over time (Atwood et al., 2005).

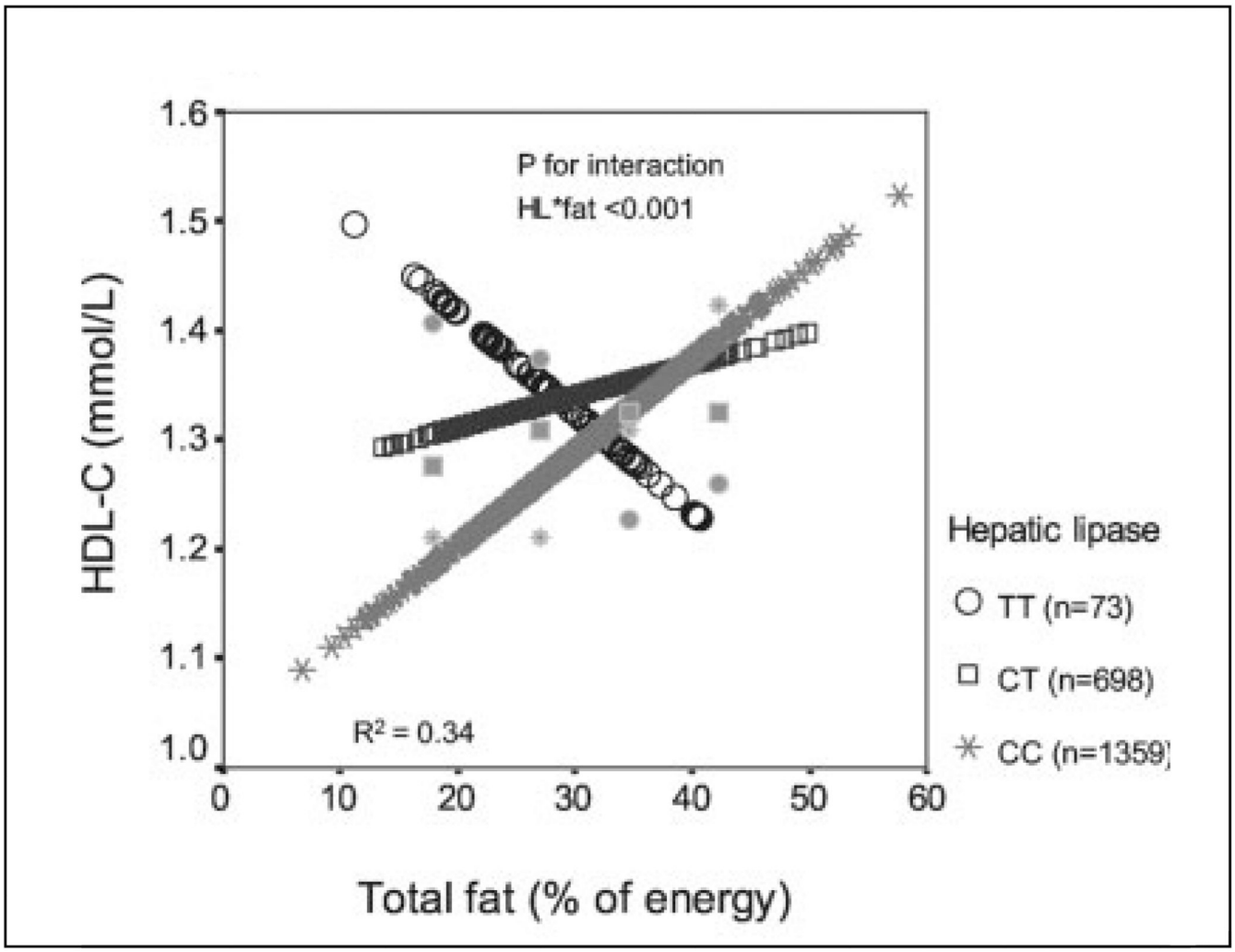


Figure 8.
Dose dependent reaction of three genotypes of the Hepatic Lipase gene in relation to HDL concentration (Ordovas et al., 2002)

Table 1

Geographic and ethnic identities of participants in the FHS population

Geographic diversity		Geographic diversity	
Birth place	***Percent***	***Ethnicity***	***Percent***
Framingham	19.15	England, Scotland and Wales	19.86
Other regions of Massachusetts	40.31	Ireland	14.95
Other regions of New England	9.79	Italy	19.00
Other US regions	9.81	French Canadian	2.26
Canada	5.46	Other Canadian	2.63
England, Scotland, Wales	1.28	Eastern Europe	5.93
Ireland	1.37	Western Europe	31. 77
Italy	7.26	Other	2.67
Other	3.32	Unknown	0.94
Unknown	2.26		
Total	100.00	Total	100.00

Table 2

Classes of phenotypic data collected on the participants of the FHS population

Data categories	Routine examinations
Physical exams	Anthropometry, blood pressure, lungs, heart, abdomen, ABI, neurological, cognition
Lifestyle and habits	Smoking, alcohol, exercise, diet, psychosocial factors
Medical history	Medications, hospitalization, diagnostic testing, cancer
Laboratory analysis	Lipids, diabetes, kidney, novel biomarkers, DNA
Non-invasive	ECG, echo, Holter monitor, carotid, vascular testing, PFT, brain and cardiac MRI, computed tomography
Endpoints	CVD, cancer, neurological, pulmonary, bone, cause-specific mortality

Table 3

Major risk factors of coronary heart disease

Modifiable	Probably modifiable	Fixed
Lipids: total cholesterol, HDL, LDL, triglycerides	Lipids: Lp(a), Oxidized LDL	Age
	Left ventricular hypertrophy	Sex
Blood pressure	Glucose intolerance	Family history
Diabetes	Hematological	
Obesity	Stress	
Sedentary lifestyle		
Alcohol intake		
Smoking		

Low density lipoprotein cholesterol (LDL); High density lipoprotein cholesterol (HDL), Lipoprotein (a) Lp(a); Wilson (1994).

Table 4

Means and standard deviations (in parenthesis) of certain variables in the FHS population among the three generation cohorts.

Cohorts	**Original Cohort 1948–1953**		**Offspring Cohort 1971–1975**		**Third Generation 2002–2005**	
Variables	**Men N=2336**	**Women N=2873**	**Men N=2483**	**Women N=2641**	**Men N=1912**	**Women N=2183**
Age, years	44 (9)	44 (9)	37 (11)	36 (10)	40 (9)	40 (9)
Current smoking, %	78	41	45	44	19	16
Systolic BP, mm Hg	136 (19)	135 (24)	126 (16)	118 (16)	121 (13)	113 (14)
Diastolic BP, mm Hg	86 (12)	84 (13)	82 (11)	76 (10)	78 (9)	73 (9)
Hypertensive medication, %	0	0	4	3	10	7
Hypertension, %	45	39	26	13	13	8
BMI, kg/m^2	25.8 (3.5)	25.4 (4.7)	26.4 (3.7)	24.0 (4.6)	27.9 (4.7)	26.0 (6.1)
BMI ≥30kg/m^2, %	12	15	15	10	26	21
Blood glucose, mg/dl	82 (24)	82 (20)	106 (16)	99 (15)	99 (18)	92 (18)
Total cholesterol, mg/dl	221 (43)	221 (46)	201 (40)	192 (39)	193 (37)	185 (34)
HDL cholesterol, mg/dl			44 (12)	56 (15)	47 (12)	61 (16)
Lipid lowering medication, %			1	0.3	11	4
Prevalent CVD, %	4	2	3	1	2	1

Adv Genet. Author manuscript; available in PMC 2011 January 3.

Table 5

Heritability estimates of some of the traits that are related to cardiovascular diseases and aging

Abdominal aortic calcification	0.49	O'Donnell et al. (2002)
Age at Natural Menopause	0.52	Murabito et al (2005)
Body mass index (BMI)	0.39	Liu et.al. (2003)
Glucose	0.23	
Systolic Blood Pressure	0.24	
High Density Cholesterol	0.40	
Total Cholesterol (TC)	0.47	
Triglycerides (TG)	0.42	
Low Density Lipoprotein(LDL)	0.50	
TG/HDL ratio	0.45	
LDL/HDL ratio	0.46	
TC/HDL ratio	0.46	
Creatinine	0.29	Fox (2004)
Estimated glomerular filtration rate	0.33	
Creatinine clearance	0.46	
Bone mineral density	0.47–0.67	Karasik et al (2003)
Hand osteoarthritis	0.28–0.34	Demissie et al. (2002)
Heart rate variability	0.13 – 0.23	Singh (1999)
Left ventricular mass	0.24–0.32	Post et al (1997)
Mean arterial pressure	0.33	Mitchell et al (2005)
Carotid femoral pulse wave velocity	0.40	
Brachial artery diameter	0.33	Benjamin (2004)
Flow-mediated dilation %	0.14	
Internal carotid intimal medial thickness	0.35	Fox et al.(2003)
Platelet aggregation	0.48–0.62	O'Donnell et al (2001)
QT interval	0.25	Newton-Cheh et al (2004)
White matter hyperintensity	0.55 (men) 0.52 (women)	Atwood et al (2005)
N-terminal proatrial natriuretic peptide brain natriuretic peptide (BNP).	0.44 0.35	Wang et al (2003)
Intercellular Adhesion Molecule-1	0.24	Keaney et al (2004)
C-reactive protein	0.28	Dupuis et al (2005)

Intercellular adhesion molecule –1	0.30	
Interleukin-6	0.14	
Monocyte chemoattractant protein –1	0.44	

Table 6

Genetic (above the diagonal) and phenotypic correlations (below diagonal) and environmental correlations (in parenthesis) among five risk factors (Martin et al., 2003)

	Cholesterol	HDL-C	SBP	TG	BMI
Cholesterol	-	− 0.06	0.04	0.32	0.11
HDL-C	0.12 (0.27)	-	0.22	−0. 46	−0.13
SBP	0.03 (0.02)	0.03 (0.13)	-	0.29	0.01
TG	0.35 (0.38)	− 0.34 (−0.24)	0.10 (0.02)	-	0.03
BMI	0.08 (0.06)	− 0.20 (−0.24)	0.16 (0.22)	0.18 (0.29)	-

Table 7

Chromosomal locations of quantitative trait loci and the associated LOD scores for various phenotypes. Only LOD scores above 3.0 have been listed.

Trait	Chromosomal location	Lod Score	Reference
Blood Pressure	17q12	4.7	Levy et. al. 2000
Body mass index	6q23–25	4.6	Atwood et al. 2002
Bone mineral density	21q22.3	3.1	Karasik et. al. 2002
Haematocrit	6q23–24	3.4	Lin et al. 2005
HDL3 cholesterol	6q24.2	4.0	Yang et. al. 2005
Hypertension	10q24.32	5.5	Guo et al. 2003
Internal carotid artery Intimal medial thickness	12q24.33	4.1	Fox et. al 2004
Monocyte chemoattractant Protein-1 (MCP-1)	1q25.1	4.3	Dupuis et al. 2005
Obesity and HDL-C	2q21.3	6.2	Arya et. al. 2003
Plasma triglyceride	6q24.3	3.1	Lin 2003
Pulmonary function	6q27	5.0	Wilk et al. 2003
Waist circumference	6q23	3.3	Fox et al 2004
Weight change	20q13.12	3.1	Fox et. Al. 2005

The Demographic Assessment for Health Literacy (DAHL): A New Tool for Estimating Associations between Health Literacy and Outcomes in National Surveys

Amresh D. Hanchate, PhD[1], Arlene S. Ash, PhD[1], Julie A. Gazmararian, PhD, MPH[2], Michael S. Wolf, PhD[3], and Michael K. Paasche-Orlow, MD, MA, MPH[1]

[1]Section of General Internal Medicine, Boston University School of Medicine, Boston, MA, USA; [2]Emory Center for Health Outcomes and Quality, Emory University, Atlanta, GA, USA; [3]Institute for Healthcare Studies and Division of General Internal Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA.

OBJECTIVE: To impute limited health literacy from commonly measured socio-demographic data and to compare it to the Short-Test of Functional Health Literacy in Adults (S-TOFHLA) for estimating the influence of limited health literacy on health status in the elderly.

METHODS: The Prudential Medicare Study assesses the S-TOFHLA score, leading to a "reference standard" classification of 25% of people with inadequate literacy; the National Health Interview Survey has no such assessment. We estimated a regression of S-TOFHLA on sex, age, years of schooling, and race/ethnicity in The Prudential Medicare Study data to derive a Demographic Assessment for Health Literacy (DAHL) score, and imputed inadequate literacy to the 25% with the lowest DAHL scores. Using regression, we then examined associations between several health status measures (including hypertension, diabetes, physical and mental SF-12) and inadequate literacy (imputed or test-based).

RESULTS: Estimates of association using imputed inadequate literacy closely approximate those obtained using S-TOFHLA-based inadequate literacy for most outcomes examined.

CONCLUSIONS: As few population surveys measure health literacy, the DAHL, a readily calculated health literacy proxy score, may be useful for expanding the scope of health literacy research in national survey data.

KEY WORDS: inadequate health literacy; health status; functional status; national surveys.
J Gen Intern Med 23(10):1561–6
DOI: 10.1007/s11606-008-0699-5

With approximately 90 million American adults estimated to lack the literacy skills needed to use the health-care system[1,2], an emerging literature has begun to describe the myriad health consequences of limited health literacy[3]. The Institute of Medicine (IOM) defines health literacy as "the degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions." Limited health literacy is an independent risk factor for worse health status, hospitalization, and mortality[3,4]. Virtually all health literacy research relies on data from specially designed surveys with in-person administration of a validated health literacy test – such as the Test of Functional Health Literacy in Adults (TOFHLA) and the Rapid Estimate of Adult Literacy in Medicine (REALM)[5–7]. To date, few surveys directly measure health literacy. Most that do are small and pertain to patients in specific clinical settings. Major national population health surveys, such as the National Health Interview Survey (NHIS), the Behavioral Risk Factor Surveillance System (BRFSS), and the National Health and Nutrition Examination Survey (NHANES), have no measure of health literacy.

Multiple reports have found high correlations between test-based health literacy measures and demographic indicators such as age, ethnicity, and years of schooling[5,7–10]. Imputed measures based on combinations of these indicators have been proposed[11,12]. Miller et al. found high construct validity by showing the correlation of their imputed measure with other indicators of health literacy[12]. Since large national health surveys do capture basic socio-demographic indicators, they can support an imputed health literacy measure. Such imputation will expand the scope of health literacy research to a much wider range of measures of health status, outcomes, and interventions.

Here, we propose an imputed measure of health literacy for community-living elderly, to be called the Demographic Assessment for Health Literacy (DAHL), and examine its comparative performance as a proxy for test-based measures in models to assess the influence of health literacy on health status. It is calculated from limited, broadly available data – sex, age, years of schooling and race/ethnicity. We develop the DAHL from Prudential Medicare Study data (1997), the largest population-based health literacy study to date. Our primary objective is to assess the performance of this imputed measure as a covariate in models of health status in the Prudential Study and in the National Health Interview Survey (1997 and 2005).

Received October 1, 2007
Revised February 26, 2008
Accepted May 16, 2008
Published online July 10, 2008

METHODS

The Prudential Medicare Study (1997)

The Prudential Medicare Study is an in-person survey administered to people, aged 65 or older, newly enrolled in the

Medicare HMO plans of Prudential Healthcare in four locations (Cleveland, Ohio; Houston, Texas; South Florida; Tampa, Florida) between December 1996 and August 1997, excluding those who did not live in the community, with severe cognitive impairment, or who were not comfortable speaking either in English or Spanish. Among 8,409 persons in the sampling frame, 57 percent could not be contacted (938 enrollees), refused to participate (3,247), or were not eligible for interview (737). An additional 227 enrollees did not cooperate during data collection. The final respondent sample was 3,260 enrollees with an effective response rate of 47 percent, using American Association of Public Opinion Research's definition no. 3, or 42.5 percent assuming all non-respondents were eligible.[13] Non-respondents were slightly older with higher educational attainment and were more likely to be white and living in zip codes with higher median income[10]. Additional details of this survey have been published previously[6,10,14–18]. Starting with this sample of 3,260, we excluded 282 subjects with history of stroke, 55 with severe cognitive impairment (Mini-Mental State Examination Score<18), and 99 with a missing value in one of the fields examined in this study[18,19], leaving a final analytic sample of 2,824.

An in-person baseline survey collected data on demographic, socioeconomic, and health characteristics and administered the Short Test of Functional Health Literacy in Adults (S-TOFHLA). This test includes a 36-item reading comprehension section and a 4-item numeracy section using materials and situations likely to be encountered by the elderly in health-care settings.

Health literacy is multifaceted, with a fundamental component being reading fluency as measured by the S-TOFHLA, one of the most commonly used instruments in health literacy research[5]. S-TOFHLA correlates well with other tests of health literacy, tests of general literacy, and several health outcomes[1,3,7]. The ability to read and take action based on health-related material may be closely related to the ability to read and act on other types of materials. We refer to the construct measured by the S-TOFHLA as "health literacy," although some authors view it as referring more narrowly to general literacy skills in the context of health care[3,8].

Following others, we summarized the S-TOFHLA responses in a composite score (range: 0 to 100) and categorized scores ≤53 (the lowest quartile in the Prudential data) as "inadequate" versus a merged category of "marginal (11%) or adequate (64%)." The "marginal" group was both small and similar in all characteristics of interest to those with "adequate" literacy[6,18,19].

The outcome measures are several self-reported indicators of health status – poor or fair health, hypertension, diabetes, Short Form Health measures (physical and mental SF-12)[20], difficulty with IADL (Instrumental Activities of Daily Living), and ADL (Activities of Daily Living). As a marker of substandard preventive care, we used a self-reported dichotomous indicator for never having been vaccinated for flu.

National Health Interview Survey (1997, 2005)

The National Health Interview Sample (NHIS) is an in-person, nationally representative annual survey covering many health-related, demographic, and socio-economic measures[21]. We selected the 1997 round to match the time of the Prudential Medicare Study, focusing on the 6,972 respondents aged 65 or older. We excluded an additional 153 due to missing information on at least one study measure, resulting in a study sample of 6,819. To assess the robustness of DAHL, we also performed parallel analyses using 5,914 analogously identified respondents from NHIS 2005.

While the NHIS does not measure health literacy, NHIS data are identical or nearly identical for four of the eight outcomes captured in the Prudential Medicare Study – self-reported general health (poor/fair), hypertension, diabetes, and difficulty with ADL. Although the NHIS instrument uses a seventh item ("difficulty getting around inside the home"), the other six ADL questions in NIHS are identical to the six ADL questions used in the Prudential Medicare Study. We judged these measures to be sufficiently comparable for our study of within-cohort associations between health literacy measures and outcomes. In contrast, we viewed the NHIS and Prudential IADL measures as too different to support useful comparisons.

Defining Components of the DAHL

Our goal was to predict health literacy well from characteristics that are commonly ascertained in national surveys. Besides the DAHL, which uses four SES indicators (age, sex, race, and education), we also examined an "education only" model, as well as models that added measures to the DAHL, including difficulty in reading and understanding forms at doctor visits, seeking help for reading forms at doctor or hospital visits, frequency of reading newspapers, and current employment status. Also, following Miller, Degenholtz et al. 2007, we explored models with interactions among education, age, and race.

Statistical Analysis

First, we used data from the Prudential Medicare Study and regression to derive the equations (scoring weights) that predict S-TOFHLA from various predictors. We then applied these weights to the values of the predictors in the NHIS data to produce an imputed literacy score (DAHL) for each Prudential and NHIS respondent. This is analogous to developing a coronary heart disease risk score with data from the Framingham Heart study (which included both risk factors and coronary heart disease outcomes) and then disseminating the scoring mechanism to predict coronary heart disease risk for people in other settings where the values of the risk factors are known, but coronary health status is not known[22]. Using the Prudential data and various cutoffs (for inadequate literacy) for each model, we estimated measures of agreement (sensitivity, specificity) and discrimination (area under the receiver operating characteristic, or C-statistic) as descriptors of each model's ability to predict S-TOFHLA-based inadequate literacy[23]. Since DAHL is derived from coefficient estimates, its precision can be gauged from the 95% confidence interval (CI) of the predicted DAHL score at the mean level of factors. The model specification with the best trade-off between simplicity and predictive ability was selected as "the" DAHL. Its associated weights were then used to calculate DAHL scores for each person in the Prudential and NHIS samples. Within each sample, the imputed literacy indicator was defined analogously, with the bottom 25 percent classified as having inadequate literacy.

The DAHL was examined for how closely its association with the various health status measures approximates the association of the S-TOFHLA with that measure in the Prudential study. We first estimated reference effects using the S-

TOFHLA-based inadequacy indicator to predict each of the eight selected health status measures in the Prudential data. Each estimate was obtained from a multivariable regression (logistic for dichotomous and linear for ordinal measures) on the test-based indicator for inadequate literacy in addition to sex, age, years of schooling, race/ethnicity, household income, and marital status.

Analogous regressions were estimated using an imputed inadequate literacy indicator, however, with only household income and marital status as additional covariates, since the DAHL is completely identified by sex, age, years of schooling, and race/ethnicity. We repeated this for all eight health status measures in the Prudential study and for the four measures in NHIS – the latter for both 1997 and 2005 data. As the NHIS survey is a multi-stage stratified sample with sampling weights differing across subgroups, we used corrective survey-adjusted weights to obtain both rates and standard errors. For the Prudential Study, we follow previous studies in treating the data as self-representative[10].

Sensitivity Analysis

We evaluated the robustness of the regression estimates to the choice of threshold score for classifying those with inadequate literacy – both for the S-TOHFLA and the DAHL measures. We chose a wide range of score thresholds (from 50 to 87), while ensuring that either side of the threshold has at least 10% of sample subjects. We obtained an estimate of the association with inadequate literacy from each cut-off and compared the corresponding estimates from the two literacy measures graphically.

While all subjects in the Prudential Medicare Study are HMO enrollees, those in NHIS include Medicare HMO as well as Fee-for-Service enrollees. As these two types of NHIS enrollees may differ in terms of health literacy, socioeconomic status, and health indicators, we also examined the differences in imputed literacy scores in these cohorts separately. All statistical analyses were performed using Stata 9.2[24]. The Institutional Review Board of Boston University School of Medicine approved the study protocol.

RESULTS

We studied 2,824 subjects from the Prudential Medicare Study and 6,819 subjects from the 1997 round of NHIS. Although the Prudential Medicare Study includes only four cities/regions, its respondents are similar to those in the NHIS sample (Table 1). The S-TOHFLA-based health literacy score ranged from 0 to 100, with a mean of 73 and standard deviation of 26 in the Prudential study,

Table 2 shows the score weights for the DAHL, calculated as the sum of the score for the reference group minus adjustments for other demographic cohorts. Specifically, white women aged 65–69 with more than 12 years of schooling (the reference group) receive a mean DAHL of 91.3, the highest possible imputed score. Less schooling, older age, and other races and ethnicities result in lower imputed health literacy. The lowest possible imputed literacy score is 15.6 among Black males 85 or older with fewer than 9 years of schooling. At the mean value of the factors determining DAHL, the predicted literacy score, which by design equals the mean of the test-based score (73.0), had a 95% confidence interval of [72.2, 73.8].

Table 1. Characteristics of the Prudential Medicare Study and National Health Interview Survey (NHIS)-Elderly Samples, 1997

	Prudential Study 1997 (N=2,824)	NHIS-Elderly 1997 (N=6,819)
Female, %	58	62
Age, %		
65–69	37	27
70–74	28	27
75–79	19	22
80–84	11	15
85	5	10
Race/ethnicity, %		
Black	11	8
White	76	86
Hispanic	12	4
Other	1	2
Marital status, %		
Married	55	42
Previously married	43	54
Never married	2	5
Years of schooling completed, %		
0–8	17	19
9–11	18	18
12 or GED	34	32
>12	30	31
Annual income, %		
Less than $10,000	20	25
$10,001 - $15,000	24	13
$15,001 - $25,000	35	29
$25,001 - $35,000	9	11
$35,001+	12	23
Test-based literacy		
S-TOHFLA health literacy score, mean (SD)	71.3 (26.8)	NA
Inadequate health literacy (S-TOHFLA ≤53), %	25	NA
Self-reported health/chronic condition		
General health - fair or poor, %	24	26
Hypertension, %	45	52
Diabetes, %	14	13
Difficulty with ADL, %	3.3	5.3
Difficulty with an IADL, %	28	NA
Never had flu vaccination, %	21	NA
Physical SF-12 score, mean (SD)	45.8 (10.9)	NA
Mental SF-12 score, mean (SD)	55.2 (8.5)	NA

Note: Estimates for NHIS 1997 are weighted to adjust for the sampling design
NA = Not available in NHIS

In the Prudential data, the S-TOHFLA and DAHL are highly correlated (ρ=0.58), and a linear regression of DAHL on S-TOHFLA gives a coefficient estimate of 0.93. We defined "inadequate health literacy" in the Prudential Study as having an S-TOHFLA score in the lowest 25th percentile (≤53) and imputed "inadequacy" for the 25% of Prudential study persons with the lowest DAHL (≤62). With these definitions, 79% of cases are correctly classified by the DAHL, sensitivity for detecting "inadequacy" is 59%, and specificity, 84%. Using a DAHL threshold of 69 increases sensitivity to 72%, but lowers specificity to 77%. The area under the receiver operating curve (the C-statistic) is 0.81 [95% CI=(0.79, 0.83)], indicating that, overall, DAHL discriminates well among people with higher and lower S-TOFHLA scores. Adding interaction terms for education, race, and age to the DAHL left the C-statistic at 0.81, while predicting inadequate literacy from only the single best measure ("education") is much less effective: sensitivity

Table 2. Obtaining the Demographic Assessment of Health Literacy (DAHL)

(The Prudential Medicare Study 1997, N=2,824)	Health Literacy Score	95% CI
Reference group:		
DAHL for White women aged 65–69 with >12 years of schooling	91.3	[89.3, 93.2]
Adjustments for other groups:		
Gender		
Male	-1.8	[-3.5, -0.27]
Age		
70–74	-5.5	[-7.5, -3.5]
75–79	-10.9	[-13.1, -8.65]
80–84	-16.2	[-18.9, -13.4]
85+	-27.8	[-31.8, -23.9]
Race/ethnicity		
Black	-15.9	[-18.5, -13.4]
Hispanic	-6.7	[-9.4, –3.9]
Other	-8.7	[-15.8, -1.7]
Years of schooling completed		
0–8	-30.2	[-32.7, -27.6]
9–11	-15.9	[-18.3, -13.6]
12 or GED	-6.2	[-8.1,-4.2]

(58%), specificity (10%), and C-statistic=0.72. Augmenting the DAHL with measures for difficulty in reading forms, seeking help in reading forms, newspaper reading frequency, and current employment status only modestly improves discrimination (C-statistic=0.83).

The performance of the imputed inadequate health literacy (i-IL) as a proxy for the S-TOHFLA-based "gold-standard" indicator (IL) to quantify associations with various measures of health status is shown in Table 3. Test-based IL was associated with poorer health for all eight health-status measures, although in one case (hypertension) it was not statistically significant at the 5% level. For each of the six dichotomous and two continuous measures of health, the 95% confidence intervals for the i-IL and IL associations in the Prudential Study overlap each other. Furthermore, for the four dichotomous outcomes that are also available in NHIS, the 95% confidence interval for the i-IL association in NHIS overlaps each of the other two confidence intervals. Several point estimates are quite similar in all three situations, for example, estimated odds ratios (ORs) for self-reported poor/fair general health were 1.77 for IL and 1.78 for i-IL in the Prudential Medicare Study, and 1.70 in NHIS. In one instance (diabetes), however, the estimated ORs varied substantially (being 1.37, 1.08, and 1.29, respectively), with the association of i-IL in the Prudential study being not significant, while the other two are significant at the 5% level. The only other large difference was observed for the Mental SF-12, where the effect estimates for inadequate literacy were -2.46 when measured using IL versus -1.27 when using i-IL in the Prudential Medicare Study. This difference may be large enough to be meaningful, although even here, the confidence intervals overlap.

These results are based on using a threshold that categorizes 25 percent of the population as having inadequate literacy; in Figure 1 we illustrate the corresponding estimates for a range of threshold scores. For all the measures except "difficulty with an ADL" we found considerable stability in the OR estimates associated with S-TOFHLA-based inadequate literacy across the entire spectrum of threshold choice from 50 to 87. This stability was matched well by the estimate based on the DAHL across most of the spectrum – except at the lowest threshold scores, possibly due to small numbers (only 14% of the Prudential study has i-IL<55). A similar pattern was found for the corresponding associations in the NHIS (1997) sample – here based only on the imputed inadequate literacy indicator (Fig. 2).

While the Prudential Medicare Study sample includes only HMO enrollees, the NHIS sample includes both HMO and Fee-for-Service enrollees. Comparing these two cohorts in the NHIS sample, we found that while the HMO enrollees were more likely to be younger, better educated, and have higher income, these differences were small, and the derived rates of inadequate literacy were not statistically significant.

DISCUSSION

This study examines the performance of an imputed measure of inadequate health literacy among elderly subjects as a proxy for test-based measures commonly used in the literature. We used the S-TOHFLA-based measure of health literacy in the Prudential

Table 3. Association of Inadequate Literacy With Self-reported Health and Chronic Conditions

	Prudential Study Sample, N=2,824						NHIS 1997, N=6,819		
	S-TOFHLA-based inadequate health literacy			DAHL inadequate health literacy			DAHL inadequate health literacy		
		95% CI			95% CI			95% CI	
	OR	Low	High	OR	Low	High	OR	Low	High
Dichotomous outcome measures									
Poor/fair general health	1.77	1.42	2.21	1.78	1.45	2.19	1.70	1.49	1.94
Hypertension	1.08	0.88	1.33	1.15	0.96	1.39	1.07	0.95	1.21
Diabetes	1.37	1.04	1.81	1.08	0.83	1.40	1.29	1.08	1.54
Difficulty with ADL	1.91	1.17	3.13	2.57	1.62	4.08	2.47	1.91	3.19
Difficulty with IADL	1.77	1.41	2.22	1.52	1.25	1.86	NA		
Flu vaccination, never	1.31	1.03	1.67	1.42	1.14	1.77	NA		
		95% CI			95% CI			95% CI	
	Coefficient	Low	High	Coefficient	Low	High	Coefficient	Low	High
Continuous outcome measures									
Physical SF12	-1.70	-2.78	-0.63	-2.34	-3.34	-1.33	NA		
Mental SF12	-2.46	-3.31	-1.62	-1.27	-2.05	-0.49	NA		

NA = Not available in NHIS

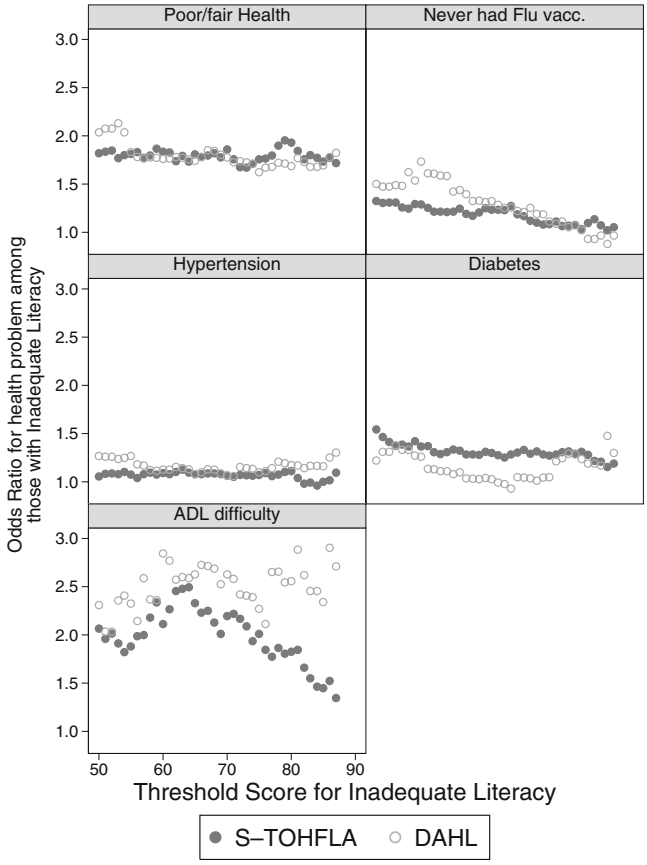


Figure 1. Association of self-reported health with inadequate literacy based on S-TOHFLA and DAHL Prudential Study 1997 (N= 2,824). Note: Each point corresponding to a threshold score denotes the effect associated with S-TOHFLA-based or DAHL-based inadequate literacy from a separate logistic regression.

Medicare Study to develop scoring weights for a parsimonious model that includes four widely available demographic indicators – sex, age, years of schooling, and race/ethnicity. Using these weights we evaluated the performance of the imputed literacy measure, the DAHL, to estimate the association with a variety of health status measures obtained in the Prudential Medicare Study (1997) as well as the elderly in the 1997 and 2005 rounds of the National Health Interview Survey (NHIS). For most of the eight health measures examined, we found similar estimates of the influence of inadequate health literacy using the imputed and test-based measures. Similarity in the estimates for the Prudential Medicare Study and NHIS is noteworthy because, while the two samples are rather similar, they differ significantly in some characteristics – for example, the NHIS elderly sample is older and less poor than the Prudential study sample.

Overall, the results support using the DAHL as a proxy for a test-based determination of inadequate health literacy in models to predict health outcomes. First, the DAHL can capture most of those who would be classified by the S-TOFHLA as having inadequate literacy. Second, even though about 20% of the sample is classified differently by the two measures, the similar magnitude and direction of associations between various health outcomes and inadequate literacy defined either way point to the underlying robustness of these associations.

The basis for the DAHL is the strong association between test-based health literacy (S-TOFHLA) and the four socioeconomic status (SES) indicators – years of schooling, age, sex, and race/ethnicity. This association is not surprising – some of these factors are causal (years of schooling, age), while others are important mediators (age, race/ethnicity, and sex). Indeed, while variation in the DAHL is dictated completely by differences in these four SES indicators, the S-TOFHLA score is obviously affected by other factors. Our findings indicate that these four SES factors capture most of the variation in S-TOFHLA, while avoiding the need for difficult to measure covariates, such as "difficulty in reading forms" that would limit a proxy measure's range of applicability.

An important implication of our sensitivity analyses is that the relationship between literacy and health outcomes appears quite stable across the range of scores. For most of the health measures examined, the odds ratio of reporting a health problem seems to be stable for much of the spectrum of both the test-based as well as imputed literacy scores. In other words, for the outcomes measured, there appear to be no particular threshold scores that identify particularly vulnerable population subgroups. Instead, the relationship appears to be linear, with potential health benefit from improved health literacy at all "levels" of literacy.

The DAHL is parsimonious in its specification compared to the imputed measure in Miller et al.[12], even though both used the same underlying socio-demographic indicators. While

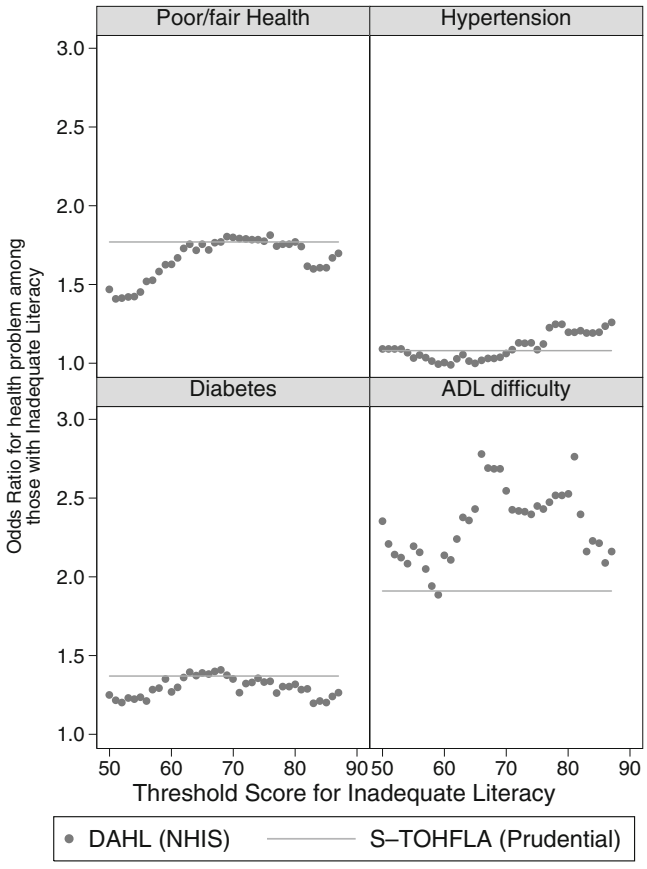


Figure 2. Association of self-reported Health with inadequate literacy based on DAHL – NHIS elderly 1997 (N=6,819). Note: Each point corresponding to a threshold score denotes the effect associated with DAHL-based inadequate literacy from a separate logistic regression. The NHIS estimate based on imputed (DAHL) inadequate literacy is compared with the S-TOFHLA-based estimate from Prudential study using the 25 percentile threshold score of 53.

the measure in Miller et al. allowed for interaction of schooling with age, Black race, and Hispanic ethnicity, the DAHL involves no interaction terms. Nevertheless, there is no loss in its discriminatory power in identifying those with inadequate health literacy, as measured by the S-TOFHLA.

Several limitations should be noted. First, the present study is limited to self-reported general health status. Analyses of other health measures should be conducted. Of the eight measures available in the Prudential Medicare Study, only four could be compared in the NHIS. Second, for health status indicators with low prevalence (such as the 3.3% prevalence for difficulty with ADL), estimates based on the imputed measure may not be stable. Future research should further evaluate this hypothesis and possibly identify a prevalence threshold that could be used as a guide for conducting analyses using the DAHL. Third, the sampling framework of the Prudential Medicare Study restricted the range of potentially important demographic characteristics that could be included in an imputed measure. For example, it is possible that including a variable for being born outside the US would improve the predictive capacity of the DAHL; however, since this variable was not collected in the Prudential Medicare Study, it could not be evaluated. There are other differences between the Prudential Medicare Study and the NHIS. First, the Prudential sample includes only Medicare HMO enrollees, while the NHIS (and other national surveys) includes both HMO and Fee-for-Service enrollees. Second, the Prudential sample includes new enrollees during an 8-month period ending August 1997; the NHIS represents the Medicare population throughout 1997. Finally, the ADL measure differs slightly across the two surveys.

To date, direct measures of health literacy require in-person evaluation, which is not done in most national health surveys. Our findings suggest that the DAHL may serve as a good proxy for estimating associations in national surveys where test-based health literacy measures are absent. Compared to the limited size and scope of the existing surveys with test-based health literacy measures, readily available national surveys, such as Medical Expenditure Panel Survey (MEPS) and Behavioral Risk Factor Surveillance System (BRFSS), offer considerably richer settings for evaluating associations of inadequate health literacy with hitherto unexamined health outcomes and utilization. These larger surveys enable examination of less common health outcomes and utilization (including, heart attacks and cardiac revascularization). In addition, longitudinal analyses on health literacy have been rare to date due to the limited availability of relevant data. Several available data sets provide the immediate opportunity to examine longitudinal hypotheses with the DAHL. Indeed, a broad range of new health literacy research questions can now be studied.

Acknowledgements: *The findings of this study were presented at the 2007 Annual Meetings of both the Society of General Internal Medicine (April 25–28, Toronto) and the AcademyHealth (June 3–5, Orlando, FL).*

Potential conflicts of interest: *None disclosed.*

Corresponding Author: *Amresh D. Hanchate, PhD; Section of General Internal Medicine, Boston University School of Medicine, 801 Massachusetts Ave, #2077, Boston, MA 01532, USA (e-mail: hanchate@bu.edu).*

REFERENCES

1. **Nielsen-Bohlman L, Panzer A, Hamlin B, Kindig D.** (National Academies Press). Health literacy: a prescription to end confusion. 2004.
2. **Kutner M, Greenberg E, Jin Y, Paulsen C.** (US Department of Education, National Center for Education Statistics). The health literacy of america's adults: results from the 2003 national assessment of adult literacy. 2006.
3. **DeWalt DA, Berkman ND, Sheridan S, Lohr KN, Pignone MP.** Literacy and health outcomes. A systematic review of the literature. J Gen Intern Med.. 2004;19(12):1228–1239.
4. **Sudore RL, Yaffe K, Satterfield S, et al.** Limited literacy and mortality in the elderly: the health, aging, and body composition study. J Gen Intern Med. 2006;21(8):806–812.
5. **Paasche-Orlow MK, Parker RM, Gazmararian JA, Nielsen-Bohlman LT, Rudd RR.** The prevalence of limited health literacy. J Gen Intern Med. 2005;20(2):175–184.
6. **Baker DW, Gazmararian JA, Williams MV, et al.** Functional health literacy and the risk of hospital admission among Medicare managed care enrollees. Am J Public Health. 2002;92(8):1278–83.
7. **Davis TC, Long SW, Jackson RH, et al.** Rapid estimate of adult literacy in medicine: a shortened screening instrument. Fam Med. 1993;25(6):391–5.
8. **DeWalt DA, Pignone MP.** Reading is fundamental: The relationship between literacy and health. Arch Intern Med. 2005;165(17):1943–1944.
9. **Parker RM, Baker DW, Williams MV, Nurss JR.** The test of functional health literacy in adults: a new instrument for measuring patients' literacy skills. J Gen Intern Med. 1995;10(10):537–41.
10. **Gazmararian JA, Baker DW, Williams MV, et al.** Health literacy among Medicare enrollees in a managed care organization. JAMA. 1999;281 (6):545–551.
11. **Paasche-Orlow M, Hanchate A.** Modeling literacy with sociodemographic characteristics and literacy activities. J Gen Intern Med. 2005;21(S4):88.
12. **Miller MJ, Degenholtz HB, Gazmararian JA, Lin CJ, Ricci EM, Sereika SM.** Identifying elderly at greatest risk of inadequate health literacy: A predictive model for population-health decision makers. Res Social Adm Pharm. 2007;3(1):70–85.
13. The American Association for Public Opinion Research. (AAPOR). Standard definitions: final dispositions of case codes and outcome rates for surveys, 4th edition. 2006.
14. **Scott TL, Gazmararian JA, Williams MV, Baker DW.** Health literacy and preventive health care use among Medicare enrollees in a managed care organization. Med Care. 2002;40(5):395–404.
15. **Howard DH, Sentell T, Gazmararian JA.** Impact of health literacy on socioeconomic and racial differences in health in an elderly population. J Gen Intern Med. 2006;21(8):857–861.
16. **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. 2006;21 (12):1215–1221.
17. **Baker DW, Gazmararian JA, Williams MV, et al.** Health literacy and use of outpatient physician services by Medicare managed care enrollees. J Gen Intern Med. 2004;19(3):215–220.
18. **Wolf MS, Gazmararian JA, Baker DW.** Health literacy and functional health status among older adults. Arch Intern Med. 2005;165(17):1946–52.
19. **Scott TL, Gazmararian JA, Williams MV, Baker DW.** Health literacy and preventive health care use among Medicare enrollees in a managed care organization. Med Care. 2002;40(5):395–404.
20. **Ware J.** The MOS 36-item short-form health survey (SF-36). In: Sederer L, Dickey B, eds. Outcomes assessment in clinical practice. Baltimore, MD: Williams & Wilkins; 1996:61–64.
21. U. S. Department of Health and Human Services. NATIONAL HEALTH INTERVIEW SURVEY, 1997 [Computer file] 2nd ICPSR version. Hyattsville, MD: US Dept. of Health and Human Services, National Center for Health Statistics [producer], 1996. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]; 2001.
22. **Brindle P, Emberson J, Lampe F, et al.** Predictive accuracy of the Framingham coronary risk score in British men: prospective cohort study. BMJ. 2003;327(7426):1267.
23. **Hanley JA, McNeil BJ.** The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 1982;143 (1):29–36.
24. StataCorp. Stata Statistical Software: Release 9 College Stataion, TX: StataCorp LP; 2005.

Exploring the Determinants of Racial and Ethnic Disparities in Total Knee Arthroplasty

Health Insurance, Income, and Assets

Amresh D. Hanchate, PhD, Yuqing Zhang, DSc,† David T. Felson, MD, MPH,† and Arlene S. Ash, PhD**

Objective: To estimate national total knee arthroplasty (TKA) rates by economic factors, and the extent to which differences in insurance coverage, income, and assets contribute to racial and ethnic disparities in TKA use.

Data Source: US longitudinal Health and Retirement Study survey data for the elderly and near-elderly (biennial rounds 1994–2004) from the Institute of Social Research, University of Michigan.

Study Design: The outcome is dichotomous, whether the respondent received first TKA in the previous 2 years. Longitudinal, random-effects logistic regression models are used to assess associations with lagged economic indicators.

Sample: Sample was 55,469 person-year observations from 18,439 persons; 663, with first TKA.

Results: Racial/ethnic disparities in TKA were more prominent among men than women. For example, relative to white women, odds ratios (ORs) were 0.94, 0.46, and 0.79, for white, black, and Hispanic men, respectively ($P < 0.05$ for black men). After adjusting for economic factors, racial/ethnic differences in TKA rates for women essentially disappeared, while the deficit for black men remained large. Among Medicare-enrolled elderly, those with supplemental insurance may be more likely to have first TKA compared with those without it, whether the supplemental coverage was private [OR: 1.27; 95% confidence interval (CI): 0.82–1.96] or Medicaid (OR: 1.18; 95% CI: 0.93–1.49). Among the near-elderly (age 47–64), compared with the privately insured, the uninsured were less likely (OR: 0.61; 95% CI: 0.40–0.92) and those with Medicaid more likely (OR: 1.53; 95% CI: 1.03–2.26) to have first TKA.

Conclusions: Limited insurance coverage and financial constraints explain some of the racial/ethnic disparities in TKA rates.

Key Words: racial and ethnic disparity, total knee arthroplasty, access constraints, insurance, out-of-pocket costs

(*Med Care* 2008;46: 481–488)

From the *Section of General Internal Medicine, Boston University School of Medicine; and †Clinical Epidemiology Research and Training Unit, Boston University School of Medicine, Boston, Massachusetts.

This study was partially funded by NIH AR47785.

Reprints: Amresh D. Hanchate, PhD, Section of General Internal Medicine, Boston University School of Medicine, 801 Massachusetts Ave, #2077, Boston, MA 01532. E-mail: hanchate@bu.edu.

ISSN: 0025-7079/08/4605-0481

Total knee arthroplasty (TKA) is increasingly common with over 431,000 procedures performed nationwide in 2004.[1] For persons with severe and potentially disabling osteoarthritis, TKA is "efficacious and cost-effective . . . [it] relieves pain and reduces functional disability."[2] As the US population ages, TKA use is expected to accelerate.[3]

Racial and ethnic disparities in TKA rates, especially among men, are striking. Sharply lower rates of TKA among elderly minorities have now been established by several recent studies using comprehensive administrative data for Medicare beneficiaries.[4–6] Skinner et al,[5] using data from virtually all TKAs performed among Medicare Fee-for-Service (FFS) enrollees during 1998–2000 (N = 431,726), found the TKA rate for black men (1.84/1000) to be only 38% of that for non-Hispanic white men (4.84); the rate for Hispanic men was intermediate (3.46). Among women, the corresponding rates were higher overall and less disparate: 5.97 for non-Hispanic whites, 5.37 for Hispanics, and 4.84 for blacks. Similar disparities in TKA have been noted in national data for over 2 decades. Despite increases in TKA use for everyone, the white-minority gap in TKA use has been growing.[7,8]

Although several studies have examined the role of patient attitudes and preferences toward major surgery,[9–12] religious beliefs[13] and willingness to use complementary and traditional care modalities,[14] little work has focused on racial and ethnic disparities in TKA use. It seems likely that racial and ethnic differences in financial constraints contribute to the TKA disparities, because the surgery is itself expensive—nationally the median inpatient cost exceeded $29,000 in 2004,[1] and there are substantial rehabilitation costs as well as the potential for lost wages. Even for those with full (parts A and B) Medicare FFS coverage, out-of-pocket expenses could reach several thousand dollars. Recent literature points to the increasing burden of out-of-pocket expenditures, even among insured populations.[15–19]

Two previous studies of the role of financial constraints in TKA disparities among Medicare FFS beneficiaries leave a confusing picture. Mahomed et al[4] concluded that those "whose income level was low enough to qualify for Medicaid supplementation were much less likely to undergo total knee replacement than individuals who did not receive Medicaid supplementation," whereas Skinner et al[6] saw "little associ-

ation between socioeconomic status and the rate of TKA." As both studies use Medicare data, part of the difference in results is due to model specification—for instance, in Mahomed et al,[4] the Medicaid recipients are compared with a reference group that includes those with and without private supplemental coverage.

A study of Canadians aged 55 or older concluded that those with less education and lower income were more likely to need TKA and similarly willing to undergo TKA as those with more education or income.[20] To the extent that these findings apply to the United States, lower TKA utilization among minorities with lower socioeconomic status (SES) is not necessarily due to unwillingness to undergo TKA.

A related study of disparities in joint (knee and hip) replacement based on a nationally representative (US) longitudinal survey sample of 6159 Medicare-enrolled adults (age 69 or older) found that those with supplementary Medigap coverage were more likely to have a joint replacement compared with those without.[21] The apparent difference with Skinner et al[6] may be due to the more detailed individual-level financial and insurance coverage measures in Dunlop et al.[21]

Although the data used in this study are from the same survey source [Health and Retirement Study (HRS)] as that used in Dunlop et al,[21] a key distinction is that we exclude hip replacements and examine TKAs exclusively. Differences by gender and race/ethnicity are also examined separately. This follows recent evidence that not only do utilization rates of knee and hip replacements vary considerably, they also differ systematically by gender and race/ethnicity.[4,6,9,21,22]

There is now much evidence of the association of insurance coverage and other economic indicators with racial and ethnic disparities.[23,24] Among Medicare enrollees, minorities are less likely to have supplemental insurance coverage, exposing them to higher out-of-pocket costs.[23,25] Also minority Medicare beneficiaries report lower rates of office visits, including those to specialists, as well as fewer diagnostic services.[18] Among the poor and previously uninsured, Medicaid coverage is associated with greater use of both preventive and curative health care services.[26,27]

METHODS

Empirical Model

Our empirical model is based on a standard economic model of individual demand for health care.[28] The direct individual cost ("price") of TKA is the out-of-pocket expenses incurred, which differs across individuals by the comprehensiveness of their insurance. Thus, given similar demographic and health conditions, the probability of receiving TKA is greater for those with more comprehensive insurance coverage (lower out-of-pocket expenditures) and more financial resources (income, savings, and other assets). The reduced form model for observing TKA from individual i in year t is specified as:

$$TKA_{it} = f(DEM_{it}, HLTH_{it}, INS_{it}, INC_{it}, UNEMP_{it}, WLTH_{it}, EDUC_{it})$$

where DEM and HLTH are demographic and health indicators, INS is health insurance coverage type, INC and WLTH are household income and wealth, and UNEMP is employment status. Highest educational achievement (EDUC) is included as a proxy for long run income earnings.

Data

We used the longitudinal HRS data from a nationally representative sample of 26,703 individuals born before 1942 and their spouses or partners. Administered by the Institute for Social Research (University of Michigan), the sample resulted from pooling (in 1998) 4 distinct age-based cohorts—2 of which had been surveyed biennially since 1992/1993, whereas the other 2 were formed in 1998.[29] HRS was designed to improve our understanding of the health dynamics of aging past age 50, including the relationship of health to economic, social, and demographic factors. Thus, it collects rich information from all these domains. The measures of income, assets, work status used here have been constructed by HRS researchers.[30]

Study Subjects

Of the 19,973 persons who completed the 1998 survey round, we excluded 1534 subjects who: (1) had a previous (pre-1998) history of knee or hip arthroplasty (n = 241), (2) belonged to none of the 3 selected racial/ethnic groups (n = 409), or (3) had incomplete covariate data (n = 884); 18,439 study subjects remained.

Analytic Data Structure

The HRS cohort was tracked over 3 biennial rounds (2000, 2002, and 2004). We excluded follow-up observations for any of the following: (1) reported TKA in previous round, (2) death, or (3) dropped from survey or incomplete covariate information. To minimize confounding from reverse causation, the outcome measure (whether or not first TKA was reported), was cross-matched with independent covariates reported in the previous survey round. For instance, "presence of a first TKA" in 2002 was matched with demographic, health, and economic conditions as reported in the 2000 survey round. There are 55,469 observations for the 18,439 study subjects—33% of subjects have 8 years of exposure (ie, 4 biennial survey rounds), 43% have 6 years of exposure, 13% have 4 years of exposure, and the remaining 11% have only 1 survey round response (2 years of exposure). The cross-matching is incomplete for some in 1998 because 3715 persons (21%) have no pre-1998 information. For this group, only outcome observations from 2000, 2002, and 2004 rounds are used in the analyses.

Outcome Measure

The main outcome measure is a binary indicator for a first TKA observed in the survey period. Those who responded affirmatively to "have you had or has a doctor told you that you have arthritis or rheumatism?" were asked: "in the last 2 years (or since previous interview) have you had surgery or any joint replacement because of arthritis?" If "yes" the follow-up query was: "which joint was that?" (1 = hip(s); 2 = knee(s); 3 = hand/wrist area; 4 = foot/ankle area; 5 = shoulder(s); 6 = spine; 7 = other). TKA was identified

by the response "knee(s)." Those with a prior knee or hip arthroplasty were excluded. Data on prior joint (knee or hip) arthroplasty was available for the 2 cohorts surveyed before 1998—these comprised 79% of the study individuals and 81% of observations. Corresponding information is not available from others—whether this has any systematic effect on identifying first TKA is unclear because those with a prior TKA in one leg are more likely to have another TKA, whereas those with previous TKA in both legs are less likely. The validity of the self-reported TKA outcome depends on recall accuracy. Most studies of the recall of major surgical procedures (hysterectomy, tubal sterilization, cancer surgeries) report high rates of agreement with medical records.[31–34]

Health Insurance, Income, and Assets

Access to health insurance was categorized in 7 groups. Four of these categories pertain to Medicare beneficiaries with differential out-of-pocket burden: Medicare FFS only, Medicare FFS with Medigap or other secondary coverage, Medicare FFS with Medicaid, and Medicare HMO. The remaining categories distinguish insurance status among those without Medicare: Medicaid, other insurance (employer-sponsored, Department of Defense health insurance [TRICARE], or Veterans Administration [VA]), and none. Household income was adjusted for inflation and household size, by dividing by the square root of household size.[35] Income was summed from all sources: earnings, capital income, employer pension, all Social Security receipts, unemployment or workers compensation, other government transfers (veterans benefits, welfare, and food stamps), and other sources (such as alimony and inheritance). When respondents were unable or unwilling to specify dollar amounts, income ranges were substituted and later used to impute those amounts.[30] Earnings were imputed for 7.5% of respondents. Assets were measured from the value of all forms of nonhousing assets, including stocks, bonds, individual retirement accounts, mutual funds, savings and checking balances, debt, vehicles, and businesses. The Consumer Price Index was used to express income and asset amounts in 1993 dollars. Regional differences in cost of living, access, and other factors were captured by dummy indicators for the regions, which is equivalent to limiting comparisons of TKA rates across persons within each region. All financial indicators were included in the multivariate regression analyses. Although they were significantly correlated, the highest correlation (between income and assets) was only 0.56. Because the data comprise a mix of retirees and employed, reliance on assets is more important for one group than for other.

Other Measures

The key demographic variables were age, gender, and race. Using the fields for race (white or black) and Hispanic ethnicity, the study sample was categorized as Hispanic, non-Hispanic black, and non-Hispanic white. Health status was captured using several indicators for selected chronic conditions and limitations on physical activities. Presence of a chronic health condition was based on the question, "Has a doctor ever told you that you have . . . .?". The conditions include high blood pressure/hypertension, diabetes, cancer, lung disease, and heart disease ("heart attack, coronary heart disease, angina, congestive heart failure, or other heart problems"). Note that although these conditions are determined from self-report of a physician diagnosis, the survey query for presence of arthritis (noted earlier) includes nonphysician diagnosed self-report of arthritis. Physical functional limitations were self-reported as difficulty with each selected activity. Following Dunlop et al,[21] binary indicators (1/0) were used to measure whether a respondent had any difficulty in performing several activities of daily living that involve lower extremities: walking 1 block, getting up from a chair, climbing 1 flight of stairs, and stooping or crouching. We also included body mass index as a health status indicator.

Statistical Methods

We tested cross-tabulations of differences in covariates by race and ethnicity with χ^2 tests. We estimated the reduced form equation of the model associating the first TKA with various indicators of potential financial constraint, health conditions, and demographics. Given the binary outcome (presence of first TKA) in a longitudinal data structure, a random-effects (longitudinal) logistic model was estimated first for the whole sample and then for the subsample of those with arthritis; this explores the extent to which racial/ethnic disparities in first TKA are accounted for by disparities in self-reported arthritis. Estimates of odds ratios (ORs) mentioned in the following sections are statistically significant at 5% level (OR not equal to 1) unless otherwise noted. All analyses were performed using STATA 9.2,[36] with sampling weights to reflect national distributions for the demographic groups. The Institutional Review Board at Boston University School of Medicine approved the study protocol.

RESULTS

The 18,439 persons in the sample were observed for an average of 6 years (Table 1); 57% were women, 8% Hispanic, and 14% non-Hispanic black. A total of 663 persons had a TKA between 1998 and 2004. The estimated crude rate of first TKA in the nation was 5.9 TKAs per 1000 persons per year [95% confidence interval (CI): 5.4–6.4]. Among women, crude rates varied modestly by race and ethnicity, being 4.2 in Hispanics, 6.4 in blacks, and 6.8 in whites, respectively. Among men, racial/ethnic TKA rates differed more, with the analogous rates being 3.0, 4.3, and 5.4.

Table 2 shows summary statistics of the independent covariates by race/ethnicity. Every covariate differed by race and ethnicity (all $P < 0.001$). Broadly, blacks and Hispanics had more illnesses and more difficulties with physical functioning than whites. Specifically, more blacks had hypertension (63%), diabetes (23%), and were overweight/obese (75%) than whites (45%, 12%, and 61%, respectively). Hispanics had higher rates of diabetes and obesity compared with whites, but lower rates of heart disease, lung disease, and cancer. A key indicator of the need for TKA is difficulty with physical activities of daily living. More blacks than whites reported difficulty with each of the 4 indicators of physical functional health of lower extremities studied; more Hispanics had difficulties than whites on 3 of these 4 indicators (Table 2). Because arthritis was also more common among blacks and Hispanics than whites (Table 1), the need for TKA seems to be greater for blacks and Hispanics than for whites.

TABLE 1. Sample Counts and Mean Rates of Total Knee Arthroplasty (TKA) by Race/Ethnicity and Sex

	Women			Men			
	White	Black	Hispanic	White	Black	Hispanic	Total
Sample							
Persons	8026	1597	800	6414	989	613	18,439
Observations	24,250	4760	2419	19,311	2883	1846	55,469
Mean exposure duration (yr)	6.0	5.8	5.9	6.0	5.7	5.8	5.9
Total no. persons with first TKA	332	58	21	219	18	15	663
Persons with self-reported arthritis (%)	68	73	67	56	56	47	63
Mean TKA rate (sample-weighted)*	6.8	6.4	4.2	5.4	3.0	4.3	5.9
95% confidence interval of mean TKA rate	6.0–7.6	4.4–8.4	1.9–6.5	4.6–6.1	1.4–4.6	1.9–6.7	5.4–6.4

*Rates per thousand persons per year. These nationally representative rates and confidence intervals are obtained by adjusting for the stratified sampling using sampling weights. Two-year recall rates are also converted to the 1-year rate presented here.

Blacks and Hispanics have fewer financial resources than whites (Table 2). For instance, only 13% of whites had annual adjusted incomes less than $10,000, fully 39% of blacks, and 46% of Hispanics had incomes below this near-poverty level. Similar patterns were observed for household assets and highest educational achievement. The figures for health insurance type are stratified by age, based on eligibility for Medicare's near-universal, aged-based insurance benefit: age 64 or younger versus 65+ (Table 2). In the younger group, 78% of whites had private, VA, or TRICARE coverage, in contrast to 60% of blacks and only 50% of Hispanics. Other coverage in the under age 65 group, that is, Medicaid or Medicare (disability), was highest for blacks (24%), whereas noninsurance was highest for Hispanics (33%).

In the 65+ group, fully 47% of whites had Medicare FFS plus supplementary coverage, as compared with only 28% of blacks and 14% of Hispanics. The rankings were reversed for Medicare plus Medicaid coverage, at 3%, 15%, and 28%, for the same groups, respectively.

Adjusted differences in TKA rates by race and ethnicity are obtained via regression (Table 3). One model adjusts only for age and illness burden (base model). A comprehensive regression additionally adjusts for economic indicators (economic model). We estimate the economic model both using the whole sample and the 57% of the sample with self-reported arthritis. In all regressions, ORs are estimated with reference to white women, aged 65 or older, because it is the group with the highest crude rates. Given the low base rates of TKA, the estimated ORs can be interpreted as risk ratios.[37] TKA rate differences adjusted for age and illness burden (base model) indicate no significant difference in TKA rate between white men and women (OR = 0.94 for white men). The base model also points to significant underutilization of TKA among blacks and Hispanics. The extent of underutilization for black and Hispanic men is much larger (OR: black = 0.46; Hispanic = 0.79) than for black and Hispanic women (OR: black = 0.72; Hispanic = 0.60). All differences except that for Hispanic men are statistically significant at the 5% level. The Hispanic sample sizes are small, with only about 600 men and 800 women.

The economic model (all sample) results indicate sizable and statistically significant associations with type of health insurance, household income, and assets. Note that the sex and race/ethnicity ORs describe differences from white women, whereas other ORs refer other contrasts. For example, among those over the age of 65, coverage that is more comprehensive than basic Medicare FFS coverage seems to be associated with higher TKA rates. This association occurs separately for those with supplemental coverage from Medicaid coverage (OR: 1.27), or from private, VA, or TRICARE insurance (OR: 1.18), as well as for those with Medicare HMO coverage (OR: 1.28). Although none of these factors is individually significant at the 0.05 level, they all point to the same conclusion that additional coverage is associated with higher TKA use. Comparing the main categories for those 47–64 years old reveals that the uninsured had a much lower TKA rate (OR: 0.61) and those with Medicaid, a much higher rate (OR: 1.53) than those with private (employer-sponsored) or TRICARE coverage.

As for associations with income, those in the lowest income category (under $10K) have an estimated OR of 0.75, and those in the next higher category ($10K–$20K) have an OR of 0.79 as compared with those in the highest income tier. Education below high school graduation is associated with 27% lower risk for TKA rate than for the college educated.

The TKA deficits for black and Hispanic women as compared with white women in the base model disappear in the all sample economic model (statistically significant ORs of 0.72 and 0.60 become insignificant ORs of 0.94 and 0.87); the deficits for black men are also reduced (from OR 0.46 to 0.56) in moving from the base to the economic model, but remain large and significant at the 5% level, whereas the OR for Hispanic men even seems to reverse direction (from OR 0.79 to 1.08), although in neither model is there enough precision to achieve statistical significance.

An alternative, but statistically equivalent, specification of the Table 3 models is to replace the 6 gender-race/ethnicity stratified groups into multiple fields to obtain a breakdown in terms of the "pure" effects of sex and race/ethnicity as well as interactions of the two. Estimates from this specification of the economic model (not reported) indicate that the pure effects of sex and race/ethnicity are dominated by the interactions. Thus, for example, the deficit for black men is far more than could be predicted from the lower rates for blacks (overall) in

TABLE 2. Sample Characteristics (Based on N = 55,469 Observations for 18,439 Persons)

	White	Black	Hispanic	All
N (observations)	43,561	7643	4265	55,469
Age distribution, yr (%)				
47–64	52	58	63	53
65–74	26	25	23	25
75–110	23	17	13	22
Specified health problems (%)				
Arthritis	56	61	52	57
Overweight or obese	61	75	70	62
Hypertension/high blood pressure	45	63	45	47
Diabetes	12	23	21	14
Cancer	12	9	6	11
Lung disease	9	7	6	9
Heart disease	21	20	14	21
Functional difficulty with (%)				
Walking 1 block	9	16	11	10
Getting up from chair	33	40	37	34
Climbing 1 flight of stairs	12	20	19	14
Stooping or crouching	38	42	38	38
Household income, adjusted for size (%)				
<$10K	13	39	46	17
$10K–$20K	24	25	25	24
$20K+	63	36	29	58
Household assets (without house) (%)				
<$5K	14	52	52	20
$5K–$20K	13	18	19	13
$20K+	74	30	29	67
Currently employed (%)	35	35	39	36
Education (%)				
Less than high school	18	43	59	23
GED or high school	38	30	21	37
At least some college	44	26	20	40
Health insurance (%)—age 64 or younger				
No insurance	14	17	33	16
Medicare FFS only	2	5	2	2
Medicare HMO only	0.6	1.5	1.4	0.8
Medicare with medicaid	0.8	3	2	1.2
Medicare with private/DoD	2	3	0.8	2
Medicaid only	2	11	10	4
Private (employer sponsored)/DoD	78	60	50	74
Health insurance (%)—age 65 or older				
No insurance	0.5	2	2	0.7
Medicare FFS only	28	33	23	29
Medicare HMO only	14	13	21	15
Medicare with medicaid	3	15	28	5
Medicare with private/DoD	47	28	14	44
Medicaid only	4	8	9	5
Private (employer sponsored)/DoD	2	2	3	2

All differences for covariates by race and ethnicity were significant (at $P < 0.001$).
In all multivariate regressions, covariate values were taken from the survey immediately prior to the reported first TKA.

comparison with whites or men (overall) in comparison with women; at the same time, Hispanic males seem to have even higher rates of TKA than Hispanic females.

Limiting the economic model to the 57% of the people with self-reported arthritis had little effect on TKA ORs for most of the predictor variables. However, the

TABLE 3. Adjusted Odds Ratios for Receipt of Total Knee Arthroplasty: Noneconomic vs. Economic Models

	Base Model (All Sample) OR (95% CI)	Economic Model (All Sample) OR (95% CI)	Economic Model (Arthritis Subsample) OR (95% CI)
Race and sex (ref: white female)			
Black, female	**0.72** (0.52–0.99)	0.94 (0.67–1.32)	0.94 (0.67–1.31)
Hispanic, female	**0.60** (0.37–0.96)	0.87 (0.54–1.43)	0.94 (0.57–1.54)
White, male	0.94 (0.81–1.10)	0.88 (0.75–1.04)	1.01 (0.86–1.19)
Black, male	**0.46** (0.28–0.78)	**0.56** (0.33–0.95)	0.65 (0.38–1.11)
Hispanic, male	0.79 (0.47–1.30)	1.08 (0.65–1.81)	1.39 (0.83–2.33)
Age (ref: 65 or older)			
47–64	**0.66** (0.57–0.77)	0.72 (0.52–1.01)	0.80 (0.61–1.06)
Having specified health problems			
Overweight or obese	**2.64** (2.18–3.20)	**2.61** (2.15–3.17)	**2.39** (1.97–2.90)
High blood pressure	1.09 (0.94–1.27)	1.11 (0.95–1.29)	1.02 (0.87–1.19)
Diabetes	**0.67** (0.54–0.82)	**0.68** (0.56–0.85)	**0.69** (0.56–0.86)
Cancer	0.85 (0.67–1.07)	0.82 (0.65–1.03)	**0.78** (0.62–0.99)
Lung disease	0.78 (0.62–1.00)	0.85 (0.67–1.08)	0.79 (0.63–1.01)
Heart disease	0.89 (0.75–1.06)	0.93 (0.78–1.10)	0.88 (0.74–1.05)
With functional difficulty in			
Walking 1 block	**1.42** (1.16–1.75)	**1.53** (1.25–1.89)	**1.55** (1.26–1.90)
Getting up from chair	**1.79** (1.50–2.13)	**1.80** (1.51–2.14)	**1.38** (1.17–1.64)
Climbing 1 flight of stairs	**1.29** (1.06–1.58)	**1.41** (1.16–1.73)	**1.40** (1.14–1.70
Stooping or crouching	**2.98** (2.46–3.61)	**3.05** (2.51–3.69)	**2.21** (1.83–2.66)
Health insurance, age 65+			
Medicare FFS only		Reference	Reference
Medicare HMO only		1.28 (0.94–1.75)	1.28 (0.94–1.74)
Medicare with medicaid		1.27 (0.82–1.96)	1.28 (0.83–1.98)
Medicare with private/DoD		1.18 (0.93–1.49)	1.18 (0.93–1.49)
Health insurance, age 64 or younger			
Private (employer sponsored)/DoD		Reference	Reference
Uninsured		**0.61** (0.40–0.92)	**0.63** (0.41–0.94)
Medicaid		**1.53** (1.03–2.26)	1.45 (0.98–2.14)
Household income			
<$10K		**0.75** (0.58–0.98)	**0.76** (0.58–0.99)
$10K–$20K		**0.79** (0.65–0.95)	**0.77** (0.64–0.93)
$20K+		Reference	Reference
Household assets			
<$5K		**0.65** (0.51–0.84)	**0.63** (0.49–0.81)
$5K–$20K		0.86 (0.69–1.08)	0.63 (0.67–1.06)
$20K+		Reference	Reference
Employment status			
Not employed		Reference	Reference
Employed		1.17 (0.95–1.44)	**1.28** (1.04–1.58)
Highest education (%)			
Less than high school		**0.73** (0.58–0.91)	**0.69** (0.55–0.87)
GED or high school		1.02 (0.86–1.20)	0.98 (0.83–1.16)
At least some college		Reference	Reference

All regressions included indicators for year (1998, 2000, 2002, and 2004) and 9 national geographic regions. The estimates for health insurance for those aged 64 or younger are based on re-running the regression now using Private/DoD as reference. Bolded OR estimates reject null hypothesis OR = 1 at 5% significance level.

deficit for black males lessens (as the OR shifts from 0.56 to 0.65) and loses its significance (with the 95% CI becoming 0.38–1.11).

DISCUSSION

Administrative data have long shown lower rates of TKA utilization for blacks and Hispanics[5,8,38] than for whites.

However, a better understanding of plausible pathways had to wait for comprehensive survey data with sizable numbers of TKA surgeries among nonwhites. Using data from a nationally representative survey (HRS) that encompasses many more TKAs than previously, we investigated the extent to which underutilization of TKA utilization among blacks and Hispanics is associated with systematic differences in illness burden and functional limitations, prohibitive out-of-pocket costs (measured by comprehensiveness of health insurance coverage), and limited family resources (measured by income and assets). The population covered is aged 47 and older.

Crude TKA rates were notably lower for blacks and Hispanics than for whites, especially among black men. After adjusting for demographic factors, illness burden and physical functional limitations, the deficits narrowed considerably, but the black male deficit remained large (OR = 0.56) and significant at the 0.05 level. Note that given the low base incidence rate of TKA, the reported ORs are approximately equal to risk ratios.[37]

We further adjusted for economic factors associated with out-of-pocket costs and the presence of resources to pay for them, specifically, type of insurance coverage, family income and assets, and highest educational achievement. As expected, TKA utilization was greater for people with more comprehensive insurance coverage, family income, assets, and education. As blacks and Hispanics have relatively less of these assets than whites, adjusting for them "explains" some of the racial and ethnic differences in TKA rates. However, there was a striking contrast in the adjusted differences between men and women. Among women the differences essentially disappear: the deficit rates of TKA for black women (as compared with white women) decreased from approximately 28% to 6%, and for Hispanic women, from 40% to 13%. But among men, although adjusting for economic factors reduced the estimated rates of TKA under use among blacks, the effect was quite modest, with the deficit decreasing from 54% to 44%. Over all these results indicate that the lesser use of TKA among minorities is strongly associated with the greater resource and cost barriers they face.

Comparing these findings with previous studies using administrative data[4,6] is complicated by the considerable differences in measures used. The lower TKA rates for Medicare beneficiaries with Medicaid coverage (lower income group) found in Mahomed et al[4] derives from a comparison with all other Medicare beneficiaries. However, these "other Medicare beneficiaries" are a diverse group, varying substantially in insurance coverage (Medigap, Medicare HMO), income, and assets. In the study of Skinner et al,[6] the only economic indicator was zip code-level income. Thus, its finding of no association with TKA could be reconciled by our findings that the effect of income differences are partly mediated by access to and comprehensiveness of insurance coverage[6]; for instance, we found that the Medicare beneficiaries with Medicaid coverage (that is, low income) have higher TKA rates than those with Medicare FFS coverage only.

As noted earlier, data for this study were drawn from the same survey source as Dunlop et al.[21] Although we use many of the same measures, our focus is different. For one thing, the outcome in Dunlop et al[21] was any joint replacement (ie, knee or hip), we have addressed TKA alone, because racial disparity issues may differ by procedure. Unlike Dunlop et al,[21] where racial/ethnic differences in surgery rates were not examined by gender, here we examine the strikingly large disparities among men in contrast to the modest differences among women. Also, by including younger people (most of whom do not have Medicare insurance) we could examine the previous finding that blacks and Hispanics were more likely to have TKAs before age 65.[39] By focusing on economic constraints, we made finer distinctions than Dunlop et al[21] in insurance coverage, to better reflect potentially large differences in out-of-pocket costs. Nevertheless the broad thrust of the findings here are largely similar to those of Dunlop et al.[21] One significant difference is that the residual racial/ethnic disparity in their study is considerably larger (53% deficit rate) even after adjusting for economic factors[21]; this may be due to: (1) combining the sexes (because the deficit for black men is much higher than for women), (2) studying an older cohort (where all have Medicare coverage), and (3) combining hip and knee replacements. In particular, relative to knee replacement, the need for hip replacement is more likely to arise from acute events (falls) at older ages when most beneficiaries have Medicare coverage.

Even after adjusting for financial indicators, sizable disparities in TKA utilization are still estimated, especially for minority men. The literature already offers rich descriptions of alternative plausible factors. In particular, several studies have identified racial and ethnic differences in patient "preferences,"[40] in patients' attitudes toward major surgery,[9–12] their religious beliefs,[13] their willingness to use complementary and traditional care modalities,[14] and their ability to handle the uncertainty of risks and benefits from surgery.[41] Some of these so-called "preferences" may be driven by socioeconomic factors, access limitations, financial and otherwise.[42]

This study has several limitations. First, although the sample size is larger than previous studies on TKA racial/ethnic disparities, some cohorts are still not large. In particular, is contains only 609 Hispanics overall, and only about 200 with a diagnosis of arthritis. This may be why some sizable differences in TKA utilization rates are not statistically significant. Second, the estimated relationships are not causal.[43] Differences in income and assets reflect not only differences in ability to pay but also differences in health behaviors and attitudes, family and social supports, and geographic location.[24] Third, although the HRS data are rich and sizable, the health status and utilization information is self-reported and clinical indicators of the need for TKA, or pain levels, are not captured. Finally, our data cannot distinguish those who are not financially constrained from those who are simply unwilling to undergo TKA.

An important implication of our findings is that, contrary to some recent evidence, lower utilization of TKA among blacks and Hispanics may be associated with insurance coverage limitations and unaffordable out-of-pocket costs. In particular, this extends to the elderly with Medicare coverage but with prohib-

itively high out-of-pocket costs. It seems clear that public assistance to lower out-of-pocket burden for the medically needy would increase TKA utilization. In states with low take-up of Medicaid among elderly poor, assistance by providers in enrolling the eligible and needy could also be effective in reducing disparities in TKA use.[44]

REFERENCES

1. HCUPnet. HCUPnet: a tool for identifying, tracking, and analyzing naitonal hospital statistics. 2006. Available at: http://hcup.ahrq.gov. Accessed October 17, 2006.
2. Hawker GA, Wright JG, Coyte PC, et al. Differences between men and women in the rate of use of hip and knee arthroplasty. *N Engl J Med.* 2000;342:1016–1022.
3. Kane R, Saleh K, Wilt T, et al. *Total Knee Replacement.* Rockville, MD: Agency for Healthcare Research and Quality; 2003.
4. Mahomed NN, Barrett J, Katz JN, et al. Epidemiology of total knee replacement in the United States Medicare population. *J Bone Joint Surg Am.* 2005;87:1222–1228.
5. Skinner J, Weinstein JN, Sporer SM, et al. Racial, ethnic, and geographic disparities in rates of knee arthroplasty among Medicare patients. *N Engl J Med.* 2003;349:1350–1359.
6. 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.
7. Escarce JJ, McGuire TG. Changes in racial differences in use of medical procedures and diagnostic tests among elderly persons: 1986–1997. *Am J Public Health.* 2004;94:1795–1799.
8. 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.
9. Hawker GA, Wright JG, Coyte PC, et al. Determining the need for hip and knee arthroplasty: the role of clinical severity and patients' preferences. *Med Care.* 2001;39:206–216.
10. Suarez-Almazor ME, Souchek J, Kelly PA, et al. Ethnic variation in knee replacement: patient preferences or uninformed disparity? *Arch Intern Med.* 2005;165:1117–1124.
11. Chang HJ, Mehta PS, Rosenberg A, et al. Concerns of patients actively contemplating total knee replacement: differences by race and gender. *Arthritis Care Res.* 2004;51:117–123.
12. Ibrahim SA, Siminoff LA, Burant CJ, et al. Differences in expectations of outcome mediate African American/white patient differences in "willingness" to consider joint replacement. *Arthritis Rheum.* 2002;46:2429–2435.
13. Ang D, Ibrahim SA, Burant C, et al. Ethnic differences in the perception of prayer and consideration of joint arthroplasty. *Med Care.* 2002;40:471–476.
14. Ibrahim SA, Siminoff LA, Burant CJ, et al. Variation in perceptions of treatment and self-care practices in elderly with osteoarthritis: a comparison between African American and white patients. *Arthritis Rheum.* 2001;45:340–345.
15. Collins SR, Davis K, Doty MM, et al. *Gaps in Health Insurance: An All-American Problem.* Washington, DC: The Commonwealth Fund; 2006.
16. Galbraith AA, Wong ST, Kim SE, et al. Out-of-pocket financial burden for low-income families with children: socioeconomic disparities and effects of insurance. *Health Serv Res.* 2005;40:1722–1736.
17. Schoen C, Doty MM, Collins SR, et al. Insured but not protected: how many adults are underinsured? *Health Aff (Millwood).* 2005;Suppl Web Exclusives:W5-289–W5-302.
18. Gornick ME. Disparities in Medicare services: potential causes, plausible explanations, and recommendations. *Health Care Financ Rev.* 2000;21:23–43.
19. Weinick RM, Byron SC, Bierman AS. Who can't pay for health care? *J Gen Intern Med.* 2005;20:504–509.
20. Hawker GA, Wright JG, Glazier RH, et al. The effect of education and income on need and willingness to undergo total joint arthroplasty. *Arthritis Rheum.* 2002;46:3331–3339.
21. Dunlop DD, Song J, Manheim LM, et al. Racial disparities in joint replacement use among older adults. *Med Care.* 2003;41:288–298.
22. Mahomed NN, Barrett JA, Katz JN, et al. Rates and outcomes of primary and revision total hip replacement in the United States medicare population. *J Bone Joint Surg Am.* 2003;85:27–32.
23. National Research Council. *Understanding Racial and Ethnic Differences in Health in Late Life: A Research Agenda.* Washington, DC: The National Academies Press; 2004.
24. Adler NE, Boyce WT, Chesney MA, et al. Socioeconomic inequalities in health. No easy solution. *JAMA.* 1993;269:3140–3145.
25. Gross DJ, Alecxih L, Gibson MJ, et al. Out-of-pocket health spending by poor and near-poor elderly Medicare beneficiaries. *Health Serv Res.* 1999;34:241–254.
26. Currie J, Gruber J. Health insurance eligibility, utilization of medical care, and child health. *Q J Econ.* 1996;111:431–466.
27. Currie J. *The Take Up of Social Benefits.* National Bureau of Economic Research, Inc; 2004.
28. Phelps C. *Health Economics.* Boston, MA: Addison Wesley; 2003.
29. Leacock C, ed. *Getting Started with the Health and Retirement Study Version 1.0.* Ann Arbor, MI: Health and Retirement Study, University of Michigan; 2006.
30. St Clair P, Bugliari D, Chien S, et al. *RAND HRS Data Documentation.* Version D. Los Angeles, CA: Rand Center for the Study of Aging; 2004.
31. Hoffmeister M, Chang-Claude J, Brenner H. Validity of self-reported endoscopies of the large bowel and implications for estimates of colorectal cancer risk. *Am J Epidemiol.* 2007;166:130–136.
32. Phillips K-A, Milne RL, Buys S, et al. Agreement between self-reported breast cancer treatment and medical records in a population-based breast cancer family registry. *J Clin Oncol.* 2005;23:4679–4686.
33. Schootman M, Jeffe DB, West MM, et al. Self-report by elderly breast cancer patients was an acceptable alternative to surveillance, epidemiology, and end results (SEER) abstract data. *J Clin Epidemiol.* 2005;58:1316–1319.
34. Green A, Purdie D, Green L, et al. Validity of self-reported hysterectomy and tubal sterilisation. The Survey of Women's Health Study Group. *Aust NZ J Public Health.* 1997;21:337–340.
35. Smeeding TM, Rainwater L, Burtless G. *United States Poverty in a Cross-National Context.* Syracuse, NY: Maxwell School of Citizenship and Public Affairs, Syracuse University; 2000.
36. StataCorp. *Stata Statistical Software.* Release 9. College Station, TX: StataCorp LP; 2005.
37. Schwartz LM, Woloshin S, Welch HG. Misunderstandings about the effects of race and sex on physicians' referrals for cardiac catheterization. *N Engl J Med.* 1999;341:279–283.
38. U.S. Department of Health and Human Services. *Healthy People 2010.* McLean, VA: U.S. Department of Health and Human Services; 2000.
39. Kulig K, Bingham CO III, Steele LL. Racial, ethnic, and geographic disparities in rates of knee arthroplasty. *N Engl J Med.* 2004;350:305–306.
40. Katz JN. Preferences, quality, and the (under)utilization of total joint arthroplasty. *Med Care.* 2001;39:203–205.
41. Clark JP, Hudak PL, Hawker GA, et al. The Moving target: a qualitative study of elderly patients' decision-making regarding total joint replacement surgery. *J Bone Joint Surg Am.* 2004;86:1366–1374.
42. Armstrong K, Hughes-Halbert C, Asch DA. Patient preferences can be misleading as explanations for racial disparities in health care. *Arch Intern Med.* 2006;166:950–954.
43. Levy H, Meltzer D. What do we really know about whether health insurance affects health? In: McLaughlin CG, ed. *Health Policy and the Uninsured: Setting the Agenda.* Washington, DC: Urban Institute; 2001:179–204.
44. GAO. *Means-Tested Programs: Information on Access can be an Important Management Tool.* Washington, DC: United States Government Account Office; 2005. GAO-05-221.

Comparison of 12 Risk Stratification Schemes to Predict Stroke in Patients With Nonvalvular Atrial Fibrillation

Stroke Risk in Atrial Fibrillation Working Group*

Background and Purpose—More than a dozen schemes for stratifying stroke risk in patients with nonvalvular atrial fibrillation have been published. Differences among these schemes lead to inconsistent stroke risk estimates for many atrial fibrillation patients, resulting in confusion among clinicians and nonuniform use of anticoagulation.

Methods—Twelve published schemes stratifying stroke risk in patients with nonvalvular atrial fibrillation are analyzed, and observed stroke rates in independent test cohorts are compared with predicted risk status.

Results—Seven schemes were based directly on event-rate analyses, whereas 5 resulted from expert consensus. Four considered only clinical features, whereas 7 schemes included echocardiographic variables. The number of variables per scheme ranged from 4 to 8 (median, 6). The most frequently included features were previous stroke/TIA (100% of schemes), patient age (83%), hypertension (83%), and diabetes (83%), and 8 additional variables were included in ≥1 schemes. Based on published test cohorts, all 8 tested schemes stratified stroke risk, but the absolute stroke rates varied widely. Observed rates for those categorized as low risk ranged from 0% to 2.3% per year and those categorized as high risk ranged from 2.5% to 7.9% per year. When applied to the same cohorts, the fractions of patients categorized by the different schemes as low risk varied from 9% to 49% and those categorized by the different schemes as high-risk varied from 11% to 77%.

Conclusions—There are substantial, clinically relevant differences among published schemes designed to stratify stroke risk in patients with atrial fibrillation. Additional research to identify an optimum scheme for primary prevention and subsequent standardization of recommendations may lead to more uniform selection of patients for anticoagulant prophylaxis. **(*Stroke*. 2008;39:1901-1910.)**

Key Words: atrial fibrillation ■ clinical prediction rules ■ risk factors ■ stroke

The absolute risk of stroke varies widely among patients with atrial fibrillation depending on patient age and other clinical features. Estimating stroke risk is a critical first step when balancing the potential benefits and risks of chronic antithrombotic therapy for stroke prevention. Multiple stroke risk stratification schemes for atrial fibrillation patients have been proposed based on various combinations of clinical and echocardiographic predictors.[1] Although there is considerable overlap, differences alter the predicted risk status of hundreds of thousands of atrial fibrillation patients,[2–4] contributing to the inconsistent use of anticoagulation.[5]

Here, we compare 12 published schemes that stratify stroke risk in patients with nonvalvular atrial fibrillation.[6–17] The key features, the distribution of atrial fibrillation patients classified into different risk strata, and the stroke rates in test cohorts are analyzed for each scheme.

Materials and Methods

Twelve stroke risk stratification schemes were selected for inclusion based on publication in peer-reviewed English language journals from 1994 to mid 2007, beginning with the landmark Atrial Fibrillation Investigators initial analysis.[6] Schemes were identified through a computerized literature search using OVID software combining the key term "atrial fibrillation" with (separately) "risk factor" and "risk stratification." Recent review articles and a recent systematic review of independent predictors of stroke in atrial fibrillation patients[1] were also canvassed. Schemes were included if they sought to predict all stroke, ischemic stroke, or a combination of stroke, systemic embolism, or TIA in patients with nonvalvular atrial fibrillation not receiving oral anticoagulation; schemes assessing stroke in patients receiving antiplatelet therapy were included. Included reports must have explicitly proposed risk strata using ≥1 clinical or echocardiographic features and must have linked the strata to recommendations for antithrombotic prophylaxis; those assessing stroke risk factors but without proposing a specific risk stratification scheme were not considered. For schemes generated by expert groups that were serially revised, only the most recent version was included. For example, only the most recent version of the American College of Cardiology/American Heart Association/European Society of Cardiology guideline was included,[11] and an earlier iteration was not considered.[18] The single exception was inclusion of both the 2001 version[13] and the 2004 revision[16] of the American College of Chest Physicians consensus statement because the earlier scheme has been tested in 2 independent cohorts. Studies reporting the performance of specific risk stratification schemes in independent

Continuing medical education (CME) credit is available for this article. Go to http://cme.ahajournals.org to take the quiz.
Received August 22, 2007; final revision received October 10, 2007; accepted October 19, 2007.
*See Appendix for Working Group participants and affiliations.
Correspondence to Robert G. Hart, MD, Department of Neurology, University of Texas Health Science Center, 7703 Floyd Curl Drive MC 7883, San Antonio, TX 78229-3900; E-mail Hartr@uthscsa.edu

Stroke **is available at http://stroke.ahajournals.org** **DOI: 10.1161/STROKEAHA.107.501825**

Table 1. Stroke Risk Stratification Schemes for Patients With Atrial Fibrillation

Study (Year of Publication)	Derivation*	Event Type	Risk Strata	Types of Variables	During Aspirin Rx
Atrial Fibrillation Investigators (1994) (6)†	Multivariate RCT 108 events	IS	High/moderate/low	Clinical	No
Stroke Prevention in Atrial Fibrillation Investigators (1995) (7)	Multivariate RCT 73 events	IS+SE	High/moderate/low	Clinical+echo	Yes
European Atrial Fibrillation Trial Study Group (1995) (8)‡	Multivariate RCT 78 events	S	High/moderate/low	Clinical+CT	No
Atrial Fibrillation Investigators (1998) (9)	Multivariate RCT 78 events	IS	Multiple	Clinical+echo	No
Stroke Prevention in Atrial Fibrillation Investigators (1999) (10)	Multivariate RCT 130 events	IS	High/moderate/low	Clinical+echo	Yes
CHADS2 (2001) (12)	Expert consensus	IS+TIA	Multiple	Clinical	Both
American College of Chest Physicians (2001) (13)	Expert consensus	NS	High/moderate/low	Clinical+echo	NS
Framingham Heart Study (2003) (15)	Multivariate ES 83 events	S	Multiple	Clinical	Both
Van Walraven et al (2003) (14)	Recursive partitioning RCT 103 events	S+TIA	High-moderate/low	Clinical	Yes
American College of Chest Physicians (2004) (16)	Expert consensus	NS	High/moderate/low	Clinical+echo	NS
Birmingham/NICE (UK) (2006) (17)	Expert consensus	IS+SE	High/moderate/low	Clinical+echo	NS
ACC/AHA/ESC Guidelines (2006) (11)	Expert consensus	IS+SE	High/moderate/low	Clinical+echo	NS

ACC/AHA/ESC indicates American College of Cardiology/American Heart Association/European Society of Cardiology; CHADS2, congestive heart failure, hypertension, age, diabetes, secondary prevention; CT, brain computed tomograms; echo, echocardiographic; ES, epidemiological study; IS, ischemic stroke; multivariate, multivariate analysis of a derivation cohort; NICE, National Institute for Clinical Excellence; NS, not specified; RCT, randomized clinical trial; Rx, treatment; S, all stroke; SE, non-CNS systemic emboli.

Expert opinion was generally based on synthesis of multiple previous studies of independent predictors (1).

*Multivariate RCT implies multivariate analysis of non-anticoagulated participants in randomized clinical trials (not that patients were randomized on the basis of risk factors).

†The original publication presents two levels of risk: high and low; subsequently most experts have extrapolated into 3 risk tiers, with patients 65 years old or older without other risk factors as moderate risk.

‡All participants had recent stroke or TIA.

populations of nonanticoagulated patients with atrial fibrillation (including those receiving antiplatelet agents) comprised the 11 test cohorts.[2,12,15,17,19–26]

The stratification schemes were applied to a stratified random sample of 1000 patients was selected from Stroke Prevention in Atrial Fibrillation III participants[22,23] to compare the relative distribution of risk strata. This sample included 40% women and a 10% prevalence of previous stroke/TIA; 2 years were added to each participant's age to increase the mean age of this cohort to 72 years, closer to that of a large outpatient atrial fibrillation population[3,26] and pooled participants in clinical trials.[14]

Results

Seven of the 12 schemes were based on event-rate analyses of stroke predictors in a derivation cohort,[6–10,14,15] whereas the remainder originated from consensus of expert panels (Table 1).[11–13,16,17] Two investigator groups published 2 schemes each. In one instance, this was because of addition of echocardiographic variables to a previously published clinical scheme,[6,9] and in the other, analysis of a separate larger cohort given aspirin modified an earlier multivariate analysis of stroke predictors in the absence of antithrombotic therapy.[7,10] Four schemes included only clinical features, whereas 7 schemes also considered transthoracic echocardiographic variables (Table 1). One scheme stratified risk among atrial fibrillation patients with recent stroke or TIA and included results of brain CT.[8]

The schemes varied substantially in complexity: the number of variables ranged[6,11,17] from 4 to 8, with a median of 6 (Table 2). The most frequent elements were previous stroke/TIA (100%), age (83%), hypertension (83%), and diabetes (83%; Table 3). Heart failure (50%), left ventricular systolic dysfunction (50%), and systolic blood pressure (42%) were the next most frequent; coronary artery disease (33%) and female sex (25%) were included in a minority of schemes (Table 3). Schemes varied in whether they used age and systolic blood pressure as continuous or ordered categorical variables and, in the latter case, different age thresholds were used (eg, 65 vs 75 years).

Stroke rates associated with specific risk strata were assessed in independent test cohorts for 8 of the schemes, whereas for the remaining 4 schemes no test cohorts were identified (Table 4). The characteristics of patients in the 11 test cohorts varied widely, from an elderly hospital discharge cohort (mean age, 81 years) with a 25% prevalence of previous stroke/TIA,[12] to a younger outpatient-based cohort (mean age, 72 years), with a 8% prevalence of previous stroke/TIA.[24,26] Two of the test cohorts were restricted to primary prevention.[2,20] Mean observation duration ranged from 1.2 years[12] to 5.3 years,[21] with a median of 2.0 years. Two schemes[7,12] were evaluated in 6 independent test cohorts,[6,13–17] 1 in 5, and the remaining 5 in 1 or 2. In some

Table 2. Summary of 12 Stroke Risk Stratification Schemes

	Derivation Cohort			
	% of Cohort	N of Events	Event Rates, %/yr (95% CI)	Independent Testing
Atrial Fibrillation Investigators (AFI) 1994 (6)				
Age <65 yr				Pearce et al (2)
No risk factors (low risk)	15	3	1.0 (0.3–3.1)	Gage et al (12)
HTN, DM, previous stroke/TIA (high risk)	17	16	4.9 (3.0–8.1)	Wang et al (15)*
Age 65–75 yr				Gage et al (20)
No risk factors (moderate risk)	20	16	4.3 (2.7–7.1)	Fang et al (26)
HTN, DM, previous stroke/TIA (high risk)	27	27	5.7 (3.9–8.3)	
Age >75 yr				
No risk factors (high risk)	11	6	3.5 (1.6–7.7)	
HTN, DM, previous stroke/TIA (high risk)	9	13	8.1 (4.7–13.9)	
Stroke Prevention in Atrial Fibrillation Investigators (SPAF) III, 1995 (7)				
High risk	48	55	5.9 (4.5–7.8)	Feinberg et al (21)
Previous thromboembolism				SPAF Investigators (22,23)
Systolic BP >160 mm Hg				Gage et al (12)
Left ventricular dysfunction†				Wang et al (15)*
Women >75 yr				Gage et al (20)
Moderate Risk	23	12	2.8 (1.7–4.7)	Fang et al (26)
HTN, no high risk features				
Low Risk	29	6	1.0 (0.5–2.3)	
No high or moderate risk features				
European Atrial Fibrillation Trial Study Group (1995)‡ (van Latum et al) (8)				
Previous stroke/TIA‡	28	30	NR	None identified
Systolic BP >160 mm Hg	21	20	NR	
Duration of AF >1 yr	57	57	NR	
≥1 infarcts on brain CT	56	51	NR	
Cardiac enlargement on CXR	24	27	NR	
High risk = ≥3 risk factors	30	NR	NR	
Moderate risk=1–2 risk factors	61	NR	NR	
Low risk=No risk factors	9	NR	NR	
Atrial Fibrillation Investigators (AFI) 1998 (9)§				
Age <65 yr				None identified
No clinical risk factors, normal LV (low risk)	15	1	0.8 (0.2–3.0)	
No clinical risk factors, abnormal LV (high risk)	1	1	9.3 (1.3–66)	
≥1 clinical risk factor, normal LV (high risk)	15	8	3.6 (1.8–7.2)	
≥1 clinical risk factor, abnormal LV (high risk)	4	5	9.7 (4.0–23)	
Age 65–75 yr				
No clinical risk factors, normal LV (moderate)	15	8	3.2 (1.6–6.5)	
No clinical risk factors, abnormal LV (high risk)	2	2	8.4 (2.1–33)	
≥1 clinical risk factor, normal LV (high risk)	27	22	4.9 (3.2–7.4)	
≥1 clinical risk factor, abnormal LV (high risk)	4	6	12 (5.3–26)	
Age >75 yr				
No clinical risk factors, normal LV (high risk)	6	0	0 (–)	
No clinical risk factors, abnormal LV (high risk)	1	1	11 (1.4–78)	
≥1 clinical risk factor, normal LV (high risk)	9	12	8.3 (4.7–14.6)	
≥1 clinical risk factor, abnormal LV (high risk)	2	4	20 (7.4–52)	

(*Continued*)

Table 2. Continued

	Derivation Cohort			
	% of Cohort	N of Events	Event Rates, %/yr (95% CI)	Independent Testing
Stroke Prevention in Atrial Fibrillation Investigators, Aspirin Cohort, 1999 (Hart et al 1999) (10)				
High risk	22	70	7.1 (5.4–9.5)	None identified
Previous stroke/TIA				
Women >75 yr old				
Men >75 yr old+HTN				
Systolic BP >160 mm Hg				
Moderate risk (no high risk features, either of)	37	43	2.6 (1.9–3.6)	
HTN				
DM				
Low risk (no high/moderate risk features)	41	17	0.9 (0.6–1.6)	
CHADS2 (2001) (Gage et al 2001) (12)				
Congestive heart failure,# hypertension,	N/A	N/A	N/A	Gage et al (12)
Age >75 yr, diabetes=1 point each; previous				Go et al (24), Fang et al
Stroke/TIA=2 points				(26), Fang et al (19)
Risk scores range from 0–6 points				Wang et al (15)*
Low risk=0				Gage et al (20)
Moderate risk=1–2				Lip et al (17)
High risk ≥3				Healey et al (25)
American College of Chest Physicians (ACCP) 2001 (Albers et al 2001) (13)				
High risk (any of)	N/A	N/A	N/A	Pearce et al (2)
Previous thromboembolism				Gage et al (20)
HTN				
HF				
LV dysfunction by echocardiography				
Age >75 yr				
≥2 moderate risk factors				
Moderate risk (any of)	N/A	N/A	N/A	
Age 65–75 yr				
DM				
Coronary artery disease				
Low risk	N/A	N/A	N/A	
No moderate or high risk features				
Framingham Risk Score (Wang et al 2003) (15)				
Age (0–10 points)				Gage et al (20)
Gender (6 points for women)				Fang et al (26)
Systolic BP (0–4 points)				
DM (5 points)				
Previous stroke/TIA (6 points)				
Risk levels estimated for low risk patients				
0–1 points	3	NR	0.0	
0–4 points	14	NR	1.1	
0–7 points	31	NR	1.5	
Van Walraven et al for the Atrial Fibrillation Investigators 2003 (14)				
Low risk if patients do not have	24	≈12	0.5 (0.2–1.0)¶	Van Walraven et al (14)
Previous stroke/TIA				Wang et al (15)*

(*Continued*)

Table 2. Continued

	Derivation Cohort			
	% of Cohort	N of Events	Event Rates, %/yr (95% CI)	Independent Testing
Treated HTN or systolic BP >140				
Angina or previous MI				
DM				
American College of Chest Physicians (ACCP) 2004 (Singer et al 2004) (16)				
High risk (any of)	N/A	N/A	N/A	Fang et al (26)
Previous thromboembolism				
HTN				
HF				
LV dysfunction by echocardiography				
Age >75 yr				
DM				
Moderate risk				
Age 65–75 yr, no high risk features	N/A	N/A	N/A	
Low risk				
Age <65 yr, no high risk features	N/A	N/A	N/A	
Birmingham/ NICE (UK) Criteria (Lip et al 2006) (17)				
High risk	N/A	N/A	N/A	Lip et al (17)
Previous thromboembolism				
Age ≥75 yr plus DM, HTN or vascular disease				
HF or abnormal LV function by echo	N/A	N/A	N/A	
Moderate risk				
Age ≥65 yr, no high risk features				
Age <75 yr plus one of DM, HTN, or vascular disease	N/A	N/A	N/A	
Low risk				
Age <65 yr with no moderate or high risk features				
ACC/AHA/ESC Guidelines 2006 (Fuster et al 2006) (11)				
High risk	N/A	N/A	N/A	None identified
Previous thromboembolism				
≥1 moderate risk feature	N/A	N/A	N/A	
Moderate risk				
Age ≥75 yr				
HF				
HTN				
DM				
LV ejection fraction ≤35% or fractional shortening <25%	N/A	N/A	N/A	
Low risk‖				
No moderate or high risk features				

AF indicates atrial fibrillation; BP, blood pressure; DM, diabetes; HF, heart; HTN, history of hypertension; LV, left ventricular; N/A, not applicable; NR, not reported; TEE, transesophageal echocardiography; ACC/AHA/ESC, American College of Cardiology/American Heart Association/European Society of Cardiology.

*Testing only the low-risk criteria with a small, uncertain number of events.

†Recent (≤3 months) clinical congestive heart failure or left ventricular fractional shortening ≤25% by M-mode echocardiography.

‡All participants had recent stroke or TIA. Event rates reported only for the combination of stroke, systemic embolism, myocardial infarction, and vascular death (not for stroke alone). Previous stroke/TIA pertains to cerebral ischemia before the qualifying event.

§Clinical risk factors are previous stroke/TIA, history of hypertension, and diabetes mellitus. Abnormal LV means moderate-to-severe systolic dysfunction by 2-dimensional echocardiocardiography.

¶For stroke events only.

‖"Less well-validated" risk factors were female sex, coronary artery disease and age 65 to 75 years. It is unclear whether patient with ≥1 of these should be categorized as moderate risk, although it is stated that antithrombotic therapy with either vitamin K antagonists or aspirin is reasonable depending on bleeding risks, ability to safely sustain anticoagulation and patient preferences.

#"Recent" heart failure, but widely applied as heart failure without time restriction in testing cohorts.

Table 3. Comparison of Features Included in 12 Risk Stratification Schemes

Study	Age, yr	HTN	DM	Previous Stroke or TIA	Female	Heart Failure	Coronary Artery Disease	Systolic BP	Abnormal LV Function
Atrial Fibrillation Investigators (1994)[6]	≥65	+	+	+					
Stroke Prevention in Atrial Fibrillation Investigators (1995)[7]	>75*	+		++	++*	++		>160	++
European Atrial Fibrillation Trial Investigators (1995)[8]†				+				>160	
Atrial Fibrillation Investigators (1998)[9]	>65	+	+	+					+
Stroke Prevention in Atrial Fibrillation Investigators (1999)[10]	>75#	+	+	++	++‡			>160	
CHADS2 (2001)[12]	≥75	+	+	++		+			
American College of Chest Physicians (2001)[13]	≥65 >75	++	+	++		++	+		++
Framingham Heart Study (2003)[15]	+§		+	+	+			+	
van Walraven et al (2003)[14]		+	+	+			+	+	
American College of Chest Physicians (2004)[16]	≥65 >75	++	++	++		++			++
Birmingham/NICE (UK) (2006)[17]	≥65	+	+	++		++	+		++
ACC/AHA/ESC Guidelines (2006)[11]¶	≥75	+	+	++	¶	+	¶		+
Overall frequency	83%	83%	83%	100%	25%	50%	25%	42%	50%

See Table 2 for specific schemes.

+ indicates included in risk stratification; ++, heavily weighted or indicative of high rather than moderate risk.

*Stroke risk was classified as high in women >75 years, but not elderly men or younger women.

†All participants had recent stroke or TIA. Three additional risk factors unique to this scheme not included in the Table: cardiomegaly on chest x-ray, ischemic stroke on brain CT, and duration of atrial fibrillation >1 year.

‡Stroke risk was classified as high in all women >75 years old and men >75 years old with hypertension, but not younger women or elderly men without hypertension.

§In the Framingham scheme, age was divided into 11 strata.

¶Age ≥65 years, being female, and coronary artery disease were stated to be "less validated or weaker risk factors."

test cohorts, echocardiographic assessment of left ventricular function was not available and clinical heart failure was substituted; in other test cohorts, a history of hypertension was substituted for measured systolic blood pressure >160 mm Hg[12,15,20,26] compromising assessment.[7]

In the test cohorts all of the schemes predicted rank order of stroke risk (Table 4). Stroke rates in patients categorized as being at low risk ranged from 0% to 2.3% per year.[15,17] For example, patients classified as being at low risk based on the CHADS2 scheme[12] had observed stroke rates ranging from 0.5% per year (95% CI, 0.3 to 0.8)[26] to 1.9% per year (95% CI, 1.2 to 3.0),[12] although TIAs were combined with stroke outcomes in the latter study. Patients classified as being at high risk had observed stroke rates varying from 2.5% per year[26] to 7.9% per year.[23]

Comparisons of different schemes in a common test cohort are limited to a handful of studies (supplemental Table I, available online at http://stroke.ahajournals.org)[2,12,15,17,20,26,28] and sometimes are compromised by substitution of some features for others (eg, heart failure for left ventricular systolic dysfunction; history of hypertension for systolic blood pressure >160 mm Hg). When compared in this fashion, the proportions of patients categorized as being at low risk varied between schemes from 12% (stroke rate 0.1% per year) to 37% (stroke rate 0.9% per year) and those as being at high risk varied between schemes from 16% (stroke rate 4.0% per year) to 80% (stroke rate 2.5% per year).[26]

In the representative cohort of atrial fibrillation patients, the mean age was 72 years, 40% were women, and prevalences of hypertension, diabetes, heart failure, systolic blood pressure >160 mm Hg, coronary artery disease, and previous stroke/TIA were 56%, 15%, 29%, 12%, 24%, and 10%, respectively. Applying each scheme to the representative cohort (Figure), the fraction of patients categorized as being at low risk ranged from 7% to 42%. Assuming 2.8 million Americans with atrial fibrillation, application of different schemes would result in up to 980 000 more or fewer patients categorized as being at low risk.

Discussion

These 12 schemes for stratifying atrial fibrillation patients according to stroke risk reflect the spectrum of choices facing clinicians. Nearly all include previous stroke/TIA, age, hypertension, and diabetes as clinical predictors of stroke. However, the fraction of patients categorized as being at low risk and high risk varies 5- to 7-fold among schemes, and this contributes to inconsistent recommendations for anticoagulation for hundreds of thousands of patients with atrial fibrillation. "The widespread nonsystematic production of guidelines" [for anticoagulant treatment in atrial fibrillation] has led to considerable variation with implications for the quality of care and clinical decision making."[29] Little has changed

Table 4. Independent Testing of 12 Stroke Risk Stratification Schemes

Study	Test Cohort (N, Mean Age, 2°PVT, Type)	High-Risk Event Rate, 95% CI (% of Cohort)	Moderate-Risk Event Rate, 95% CI (% of Cohort)	Low-Risk Event Rate, 95% CI (% of Cohort)
Atrial Fibrillation Investigators (1994)[6]				
Gage et al[12]*	1733, 81 yr, 25%, HDC	5.4 (4.2–6.5); NR	2.2 (1.1–3.5); NR	NR
Gage et al[20]†	2014, ≈72 yr 0%, RCT	3.5 (2.7–4.5); 50%	1.7 (1.1–2.5); 39%	0.9 (0.3–2.3); 12%
Wang et al[15]‡	705, 75 yr, 14%, ES	NR	NR	0.9 (NR); 6%
Pearce et al[2]	1073, 68 yr, 0%, RCT	NR	NR	0.3 (0.0–2.3); 15%
Fang et al[26]	5588, 72 yr, 8%, PC	2.5 (NR); 62%	2.1 (NR); 25%	0.2 (NR); 13%
Stroke Prevention in Atrial Fibrillation Investigators (1995)[7]				*
Gage et al[12]	1733, 81 yr, 25%, HDC	5.7 (4.4–7.0); NR	3.3 (1.7–5.2); NR	1.5 (0.5–2.8); NR
Gage et al[20]†	2014, ≈72 yr, 0%, RCT	3.6 (2.7–4.7); 44%	2.7 (1.8–4.0); 23%	1.1 (0.7–1.8); 33%
Wang et al[15]‡	705, 75 yr, 14%, ES	NR	NR	2.3 (NR)†; 17%
SPAF Investigators[22,23]	1413, 69 yr, 15%, RCT	7.9 (5.9–10.6); 37%	3.6 (2.5–5.2); 29%	1.1 (0.6–2.0); 34%
Feinberg et al[21]	259, 74 yr, 7%, ES	3.7 (2.1–5.8); 45%	2.0 (0.7–4.7); 24%	1.7 (0.6–3.8); 31%
Fang et al[26]#	5588, 72 yr, 8%, PC	3.2 (NR); 44%	1.7 (NR); 29%	0.9 (NR); 28%
European Atrial Fibrillation Trial Investigators (1995)[8]	None identified	…	…	…
Atrial Fibrillation Investigators (1998) (9)	None identified	…	…	…
Stroke Prevention in Atrial Fibrillation Investigators (1999)[10]	None identified	…	…	…
CHADS2 (2001)[12]		3–6 points	1–2 points§	0 points
Gage et al[12]*	1733, 81 yr, 25%, HDC	7.6 (NR); 36%	3.4 (NR); 57%	1.9 (1.2–3.0); 7%
Gage et al[20]†	2014,≈72 yr, 0%, RCT	5.3 (3.3–8.4); 11%	2.7 (2.2–3.4); 66%	0.8 (0.4–1.7); 23%
Go et al,[24]	5089, 71 yr, 4%, PC	5.6 (NR); 20%	2.1 (NR); 57%	0.5 (0.3–0.8); 22%
Wang et al[15]‡	705, 75 yr, 14%, ES	NR	NR	1.7% (NR); 10%
Lip et al[17]¶	994, 69 yr, 13%, RCT	7.0 (3.9–11); 18%	2.7 (1.8–3.9); 56%	0.7 (0.1–1.6); 26%
Healey et al[25]‖	3335,70 yr, 15%, RCT	3.6 (NR); 27%	1.6(NR); 70%	NA; NA
Nieuwlaat et al[5]	4564, ≈70 yr, NR, PC	NR (NR); 22%	NR (NR); 60%	NR (NR); 18%
American College of Chest Physicians (2001)[11]				
Gage et al[20]†	2014,≈72 yr, 0%, RCT	3.0 (2.5–3.8); 77%	1.0 (0.4–2.2); 15%	0.5 (0.1–2.2); 9%
Pearce et al[2]†	1073, 68 yr, 0%, RCT	3.5 (2.6–4.7); 66%	1.2 (0.5–2.8); 20%	0.3 (0.1–2.5); 14%
Nieuwlaat et al[5]	3580,≈70 yr, ≈13%, PC	NR (NR); 82%	NR (NR); 7%	NR (NR); 11%
Framingham Heart Study (2003)[15]				
Gage et al[20]†	2014,≈72 yr, 0%, RCT	4.2 (2.8–6.1); 16%	3.2 (2.4–4.3); 35%	1.4 (1.0–2.1); 49%
Fang et al[26]	5588, 72 yr, 8%, PC	4.0 (NR); 16%	2.5 (NR); 47%	0.9 (NR); 37%

(*Continued*)

Table 4. Continued

Study	Test Cohort (N, Mean Age, 2°PVT Type)	High-Risk Event Rate, 95% CI (% of Cohort)	Moderate-Risk Event Rate, 95% CI (% of Cohort)	Low-Risk Event Rate, 95% CI, (% of Cohort)
van Walraven et al (2003)[14]				
van Walraven et al[14]	840, 70 yr, 3%, RCT	NR	NR	1.1 (NR); ≈22%**
Wang et al[15]‡	705, 75 yr, 14%, ES	NR	NR	1.9 (NR); 16%
American College of Chest Physicians (2004)[16]				
Fang et al[26]	5588, 72 yr, 8%, PC	2.5 (NR); 80%	0.9 (NR); 8%	0.1 (NR); 12%
Birmingham/NICE (UK) (2006)[17]				
Lip et al[17]¶	994, 69 yr, 13%, RCT	5.8 (3.7–8.3); 34%	2.0 (1.2–2.9); 55%	0 (–); 11%
ACC/AHA/ESC Guidelines (2006)[11]	None identified	...	...	...

2°PVT indicates fraction with previous stroke/TIA; ES, epidemiological study; HDC, hospital discharge cohort; NA, not applicable; NR, not reported; PC, prospective cohort; RCT, randomized clinical trial; SPAF, Stroke Prevention in Atrial Fibrillation.

*24% (23/94) of outcome events were TIAs, so stroke rates were ≈ 25% lower than the rates provided in the Table. LV fractional shortening and systolic blood pressure (2 high-risk features in the SPAF 1995 scheme) were not available in the data set; hence, participants were classified based on incomplete information that would shift high-risk patients into the moderate risk category.

†High risk rates are for primary prevention (i.e. excluding patients with previous stroke/TIA); for Gage et al(20) echocardiographic data were not available, and hence the variable "abnormal left ventricular systolic function" could not be included from the Stroke Prevention in Atrial Fibrillation Investigators (1995) scheme and the American College of Chest Physicians (2001) scheme; "recent heart failure" could not be assessed, and heart failure was used.

‡Point estimate reported without CI based on a small number of stroke events for AFI 1994 and CHADS2; echocardiographic LV dysfunction was not considered in the SPAF 1995 scheme.

§Rates/frequencies for CHADS2 score=1 are: Go et al (24) 1.5%/yr (1.2–1.9) including ischemic strokes and systemic emboli per 32%; Gage et al (12) 2.8%/yr (2.0–3.8) including strokes and TIAs per 27%; Gage et al (20) 2.2%/yr (1.6–3.1) per 37%; Healey et al (25) 1.2%/yr (NR) per 36%. In a UK outpatient cohort of 234 atrial fibrillation patients undergoing echocardiography, the frequency of CHADS2=0 was 24% (27).

¶Approximately half of test cohort overlapped the test cohort of Gage et al (20).

||All received clopidogrel and aspirin. CHADS2=0 were excluded unless peripheral vascular disease was present.

**Reported event rate for the validation cohort differs between abstract (1.1) and text (0.9); CI difficult to estimate precisely from figure 4.

#Systolic blood pressures were not available to identify high-risk patients.

since this statement appeared a decade ago, and in the meantime additional schemes and guidelines have proliferated.

Authorities on clinical prediction rules advocate independent testing before their general clinical application.[30–32] Several schemes have not been tested to characterize their predictive accuracy and hence cannot be compared, directly or indirectly, to others. The duration of follow-up in most derivation and validation cohorts averages 1 to 2 years, and the enduring predictive value of risk stratification schemes for longer periods is often unknown, requiring periodic reassessment of risk. The contribution of individual variables to risk stratification schemes has not been well-defined. For example, heart failure appeared in half the schemes, but this

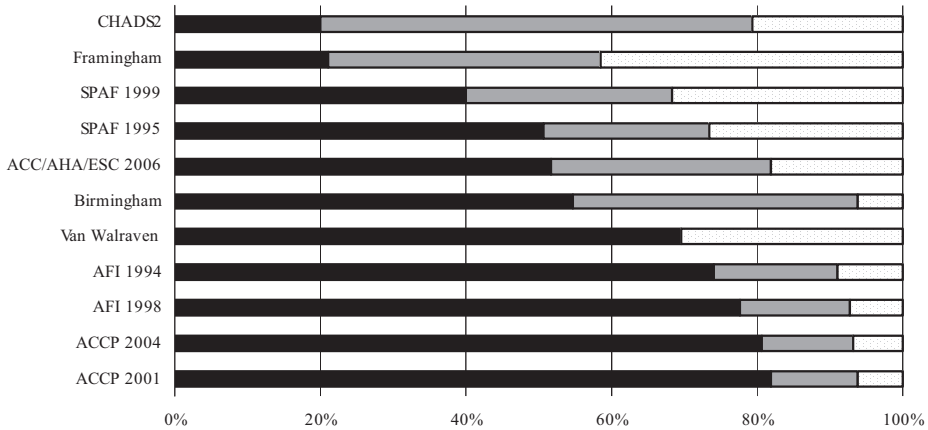


Figure. Relative distribution of patients predicted to have high (black), moderate (gray), and low (white) stroke risk by applying different risk stratification schemes to a representative cohort of atrial fibrillation patients. The mean age was 72 years and the frequencies of female sex, hypertension, diabetes, heart failure, systolic blood pressure >160 mm Hg, coronary artery disease, and previous stroke/TIA were 40%, 56%, 15%, 29%, 12%, 24%, and 10%, respectively. For the Framingham Heart Study criteria,[15] high risk was considered ≥14 points, moderate risk was 8 to 13 points, and low risk was ≤7 points. For Van Walraven et al,[14] there are only 2 risk strata: low risk and combined moderate and high risk. Because the European Atrial Fibrillation Trial criteria were intended to apply to patients with recent previous stroke/TIA, it is not included.[8] See Table 2 for study abbreviations.

clinical feature has not been validated as an independent predictor of stroke in atrial fibrillation patients.[1] Criteria used for diagnosis of heart failure have not been uniform in these studies, and the contribution of this variable to risk stratification is, therefore, unclear. The stroke risk attributable to hypertension in atrial fibrillation patients is likely to vary depending on its severity and treatment,[33] confounding application of this prevalent risk factor. Previous stroke or TIA is the most powerful risk factor[1] and, by itself, drives the successful identification of high-risk patients, regardless of the presence of other risk factors in all except 2 schemes.[12,15] The predictive value of these schemes for primary prevention (ie, for patients without previous stroke or TIA) is a more important, albeit more difficult, problem.[20]

Stroke rates in recent clinical trials[34–37] involving atrial fibrillation patients appear lower than in clinical trials completed 15 years ago.[6] Better control of blood pressure may contribute to lower stroke rates among patients with a history of hypertension,[38,39] because even modest blood pressure lowering has a substantial favorable impact on the risk of vascular events.[33] Whether absolute stroke rates among those stratified as being at high risk by any scheme are lower now than they were 10 to 15 years ago is uncertain.[38,40] In short, secular trends in stroke rates among atrial fibrillation patients may confound accurate risk prediction.

At the core of existing schemes are 4 features that have been independently and consistently associated with stroke in atrial fibrillation patients: previous stroke or TIA, hypertension, advanced age, and diabetes.[1] Other risk factors included in several schemes (eg, coronary artery disease, heart failure, female sex) have not been validated as consistent independent predictors of stroke in atrial fibrillation patients.[1,19,25] Additional possible independent predictors that are not included in current schemes (eg, estrogen replacement therapy associated with higher stroke risk, regular alcohol consumption with reduced stroke risk) have been identified,[10] but these have not been sufficiently investigated to justify application in clinical practice. The additional discriminatory power of biomarkers of thrombosis and inflammation are an area of active research.[17]

Comparison of the predictive power of available schemes with subsequent stroke in a single cohort of atrial fibrillation patients of adequate size and with a full range of variables is not currently available, and the optimal risk stratification scheme cannot be determined from existing data. The proportion of patients categorized as being at low, moderate, or high risk by a scheme will vary depending on the composition of the patient cohort to which it is applied, ie, the proportions of primary versus secondary prevention cases, proportions of elderly patients with multiple risk factors versus younger individuals with few risk factors, and the availability of echocardiographic data. Considering the inherent difficulty in distinguishing patients with stroke risk of 1% per year versus 4% per year (a determining difference regarding recommendations to anticoagulate in most guidelines), it is surprising, perhaps, that the existing schemes appear able to do so, albeit with differing results at the individual patient level.

We surveyed a scattered and complex literature on stroke risk stratification for patients with atrial fibrillation to bring its strengths and limitations into focus. We do not address the threshold of absolute stroke risk for which anticoagulation is warranted, because this depends on additional considerations, including estimated bleeding risk during anticoagulation,[41] access to quality anticoagulation monitoring, and patient preferences and values.[42] Several million people with atrial fibrillation now receive chronic anticoagulation to prevent stroke. Additional research to identify more discriminating and accurate risk models around which standard recommendations could be developed would encourage more uniform use of antithrombotic agents and would likely lead to better patient outcomes.

Appendix

Stroke Risk in Atrial Fibrillation Work Group

Robert G. Hart, MD (San Antonio, TX, USA); Lesly A. Pearce, MS (Minot, ND, USA); Jonathan L. Halperin, MD (New York, NY, USA); Elaine M. Hylek, MD, MPH (Boston, MA, USA); Gregory W. Albers, MD (Palo Alto, CA, USA); David C. Anderson, MD (Minneapolis, MN); Stuart J. Connolly, MD (Hamilton, ON, Canada); Gary H. Friday, MD (Bryn Mawr, PA, USA); Brian F. Gage, MD (St. Louis, MO, USA); Alan S. Go, MD (Oakland, CA, USA); Larry B. Goldstein, MD (Durham, NC, USA); Gary Gronseth, MD (Kansas City, KN, USA); Gregory Y.H. Lip, MD (Birmingham, UK); David G. Sherman, MD (San Antonio, TX, USA); Daniel E. Singer, MD (Boston, MA, USA); Carl van Walraven, MD, MSc (Ottawa, ON, Canada).

Disclosures

Dr Hart has served on data monitoring committees of clinical trials involving patients with atrial fibrillation sponsored by Astellas Pharmaceuticals, Sanofi-Aventis Pharmaceuticals, and Biotroniks, Inc.

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. Pearce LA, Hart RG, Halperin JL. Stratifying stroke risk in atrial fibrillation: Assessment of three clinical schemes for primary prevention. *Am J Med.* 2000;109:45–51.
3. Go AS, Hylek EM, Phillips KA, Borowsky LH, Henault LE, Chang Y, Selby JV, Singer JE. Implications of stroke risk criteria on the anticoagulation decision in nonvalvular atrial fibrillation: the Antcoagulation and Risk Factors in Atrial Fibrillation (ATRIA) study. *Circulation.* 2000;102:11–13.
4. Laguna P, Martin A, Del Arco C, Millan I, Gargantilla P on behalf of the investigators in the Spanish Atrial Fibrillation in Emergency Medicine Study Group. Differences among clinical classification schemes for predicting stroke in atrial fibrillation: Implications for therapy in daily practice. *Academic Emerg Med.* 2005;12:828–834.
5. Nieuwlaat R, Capucci A, Lip GYH, Olsson SB, Prins MH, Nieman FH, Lopez-Sendon J, Vardas PE, Aliot E, Santini M and Crijns HJGM on behalf of the Euro Heart Survey Investigators. Antithrombotic treatment in real-life atrial fibrillation patients: a report from the Euro Heart Study Survey on Atrial Fibrillation. *Eur Heart J.* 2006;27:3018–3028.
6. Atrial Fibrillation Investigators. 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;154:1449–1457.
7. 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;5:147–157.
8. van Latum JC, Koudstaal PJ, Venables GS, van Gijn J, Kappelle LJ. Algra A for the European Atrial Fibrillation Trial (EAFT) Study Group. Predictors of major vascular events in patients with a transient ischemic attack or minor ischemic stroke with nonrheumatic atrial fibrillation. *Stroke.* 1995;16:801–806.

9. Atrial Fibrillation Investigators. Echocardiographic predictors of stroke in patients with atrial fibrillation. A prospective study of 1,066 patients from three clinical trials. *Arch Intern Med.* 1998;158:1316–1320.
10. Hart RG, Pearce LA, McBride R, Rothbart RM, Asinger RW. Factors associated with ischemic stroke during aspirin therapy in atrial fibrillation: analysis of 2012 participants in the SPAF I-III clinical trials. The Stroke Prevention in Atrial Fibrillation (SPAF) Investigators. *Stroke*. 1999;30:1223–1229.
11. Fuster V, Ryden LE, Cannom DS, Crijns HJ, Curtis AB, Ellenbogen KA, Halperin JL, Le Heuzey J-Y, Kay GN, Lowe JE, Olsson SB, Prystowsky EN, Tamargo JL, Wann S. ACC/AHA/ESC 2006 guidelines for the management of patients with atrial fibrillation – Executive summary. *Circulation*. 2006;114:e257–e354.
12. Gage BF, Waterman AD, Shannon W, Boechler M, Rich MW, Radford MJ. Validation of clinical classification schemes for predicting stroke: results of the National Registry of Atrial Fibrillation. *JAMA*. 2001;285:2864–2870.
13. Albers GW, Dalen JE, Laupacis A, Manning WJ, Petersen P, Singer DE. Antithrombotic therapy in atrial fibrillation. The Sixth ACCP Conference on Antithrombotic and Thrombolytic Therapy. *Chest*. 2001;119:194S–206S.
14. van Walraven C, Hart RG, Wells GA, Petersen P, Koudstaal PJ, Gullov AL, Hellemons BSP, Koefed BG, Laupacis A. A clinical prediction rule to identify patients with atrial fibrillation and a low risk for stroke while taking aspirin. *Arch Intern Med*. 2003;163:936–943.
15. Wang TJ, Massaro JM, Levy D, Vasan RV, Wolf PA, D'Agostino RB, Larson MG, Kannel WB, Benjamin EJ. A risk score for predicting stroke or death in individuals with new-onset atrial fibrillation in the community. The Framingham Study. *JAMA*. 2003;290:1049–1056.
16. 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;126:429S–456S.
17. Lip GYH, Lane D, van Walraven C, Hart RG. Additive role of plasma von Willebrand Factor levels to clinical factors for risk stratification in patients with atrial fibrillation. *Stroke*. 2006;37:2294–2300.
18. Fuster V, Ryden LE, Asinger RW, Cannom DS, Crijns HJ, Frye RL, Halperin JL, Kay GN, Klein WW, Levy S, McNamara RL, Prystowsky EN, Wann LS, Wyse DG. ACC/AHA/ESC guidelines for the management of patients with atrial fibrillation: Executive summary. *J Am Coll Cardiol*. 2001;38:1231–1265.
19. Fang MC, Singer DE, Chang Y, Hylek EM, Henault LE, Jensvold NG, Go AS. 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;112:1687–1691.
20. Gage BF, van Walraven C, Pearce L, Hart RG, Koudstaal PJ, Boode BSP, Petersen PJ. Selecting patients with atrial fibrillation for anticoagulation. Stroke risk stratification in patients taking aspirin. *Circulation*. 2004;110:2287–2292.
21. Feinberg WM, Kronmal RA, Newman AB, Kraut MA, Bovill EG, Cooper L, Hart RG. Stroke and atrial fibrillation in a population-based cohort. The Cardiovascular Health Study. *J Gen Intern Med*. 1999;14:56–59.
22. Stroke Prevention in Atrial Fibrillation Investigators. Adjusted-dose warfarin versus low-intensity, fixed-dose warfarin plus aspirin for high-risk patients with atrial fibrillation. The Stroke Prevention in Atrial Fibrillation III randomised clinical trial. *Lancet.* 1996;348:633–638.
23. Stroke Prevention in Atrial Fibrillation Investigators. Patients with nonvalvular atrial fibrillation at low-risk of stroke during treatment with aspirin: The Stroke Prevention in Atrial Fibrillation III Study. *JAMA*. 1998;279:1273–1277.
24. Go AS, Hylek EM, Chang Y, Phillips KA, Henault LE, Capra AM, Jensvold NG, Selby JV, Singer DE. Anticoagulation therapy for stroke prevention in atrial fibrillation. How well do randomized trials translate into clinical practice? *JAMA*. 2003;290:2685–2692.
25. Healey JS, Hart RG, Pogue J, Pfeffer MA, Hohnloser SH, De Caterina R, Flaker G, Yusuf S, Connolly SJ. 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;39:1482–1486.
26. Fang MC, Go AS, Chang Y, Borowsky L, Pomernacki NK, Singer DE; for the 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.
27. Kalra L, Perez I, Melbourn A. Risk assessment and anticoagulation for primary stroke prevention in atrial fibrillation. *Stroke*. 1999;30:1218–1222.
28. Sudlow M, Thomson R, Thwaites B, Rodgers H, Kenny RA. Prevalence of atrial fibrillation and eligibility for anticoagulants in the community. *Lancet*. 1998;352:1167–1171.
29. Thomson R, McElroy H, Sudlow M. Guidelines on anticoagulant treatment in atrial fibrillation in Great Britain: variation in content and implications for treatment. *BMJ*. 1998;316:509–513.
30. Laupacis A, Sekar N, Stiell IG. Clinical prediction rules. A review and suggested modifications of methodological standards. *JAMA*. 1997;277:488–494.
31. Wasson JH, Sox HC, Neff RK, Goldman L. Clinical prediction rules. Applications and methodological standards. *N Engl J Med*. 1985;313:793–799.
32. Altman DG, Royston P. What do we mean by validating a prognostic model? *Stat Med*. 2000;19:453–473.
33. Arima H, Hart RG, Colman S, Chalmers J, Anderson C, Rodgers A, Woodward M, Neal B. Perindopril-based blood pressure lowering reduces major vascular events in patients with atrial fibrillation and prior stroke or TIA. *Stroke*. 2005;36:2164–2169.
34. ACTIVE Writing Committee on behalf of the ACTIVE Investigators. Clopidogrel plus aspirin versus oral anticoagulation for atrial fibrillation. Atrial fibrillation Clopidogrel Trial with Irbesartan for prevention of Vascular Events (ACTIVE-W). *Lancet.* 2006;367:1903–1912.
35. Sherman DG, Kim SG, Boop BS, Corley SD, DiMarco JP, Hart RG, Haywood LJ, Hoyte K, Kaufman ES, Kim MH, Nasco E, Waldo AL, and the NHLBI AFFIRM Investigators. The occurrence and characteristics of stroke events in the AFFIRM Study. *Arch Intern Med.* 2005;165:1185–1191.
36. Olsson SB, Executive Steering Committee on behalf of the SPORTIF III Investigators. Stroke prevention with the oral direct thrombin inhibitor ximegalatran compared with warfarin in patients with non-valvular atrial fibrillation: randomized controlled trial. *Lancet.* 2003;362:1691–1698.
37. SPORTIF Executive Steering Committee for the SPORTIF V Investigators. Ximelagatran vs. warfarin for stroke prevention in patients with nonvalvular atrial fibrillation. *JAMA.* 2005;293:690–698.
38. Miyasaka Y, Barnes ME, Gersh BJ, Cha SS, Seward JB, Bailey KR, Iwasaka T, Tsang TSM. Time trends of ischemic stroke incidence and mortality in patients diagnosed with first atrial fibrillation in 1980 to 2000. Report of a community-based study. *Stroke*. 2005;36:2362–2366.
39. Lip GYH, Frison L, Grind M on behalf of the SPORTIF Investigators. Effect of hypertension on anticoagulated patients with atrial fibrillation. *Eur Heart J.* 2007;28:752–759.
40. Lakshiminaryan K, Solid C, Anderson DC, Herzog CA. Validation of the CHADS2 stroke risk classification in Medicare data. *Stroke*. 2006;37:671 (abstract).
41. Gage BF, Yan Y, Milligan PE, Waterman AD, Culverhouse R, Rich MW, Radford MJ. Clinical classification schemes for predicting hemorrhage: Results from the National Registry of Atrial Fibrillation (NRAF). *Am Heart J.* 2006;151:713–719.
42. Man-Son-Hing M, Gage BF, Montgomery AH, Howitt A, Thomson R, Devereaux PJ, Protheroe J, Fahey T, Armstrong D, Laupacis A. Preference-based antithrombotic therapy in atrial fibrillation: Implications for clinical decision making. *Med Decis Making*. 2005;25:548–559.

The implications of health literacy on patient–provider communication

L Kari Hironaka,[1] Michael K Paasche-Orlow[2]

[1] Department of Pediatrics, Boston University School of Medicine, Boston, Massachusetts, USA; [2] Department of Medicine, Boston University School of Medicine, Boston, Massachusetts, USA

Correspondence to: Dr L Kari Hironaka, MD MPH, Department of Pediatrics, Boston University School of Medicine/Boston Medical Center; 91 East Concord Street, 5th Floor, Boston, MA 02118, USA; lynn.hironaka@bmc.org

Accepted 26 September 2007
Published Online First 4 October 2007

ABSTRACT

Limited health literacy has been associated with a range of adverse health outcomes including decreased use of preventive health services, poorer disease-specific outcomes for certain chronic conditions and increased risk of hospitalisation and mortality. Although the majority of research has been conducted in the adult population, there is a small and growing body of research on this subject in the paediatric literature. In this article, we will review the research on health literacy, consider the range of other communication skills associated with limited health literacy and explore strategies to improve patient–provider communication for clinicians who care for families with limited health-literacy skills.

A 4-year-old boy with asthma arrives at his pediatrician's office with significant shortness of breath. His mother is clearly anxious and frustrated. Despite giving him medicine every few hours, his breathing seems only to have worsened. When asked to demonstrate how she has been giving the medicine, his mother produces a corticosteroid inhaler. Unable to read the prescription, she had accidentally switched his "controller" medicine for his "rescue" medicine.

Over the past decade, there has been growing interest in the relationship between literacy and health outcomes. From this work, the concept of health literacy has emerged. Health literacy is "the degree to which individuals have the capacity to obtain, process and understand basic health information and services needed to make appropriate health decisions".[1] In 2004, the Institute of Medicine (IOM) published a report that summarised the health-literacy field and proposed areas of further investigation. The IOM concluded that adults with limited health literacy have less knowledge of disease self-management and health-promoting behaviours, report poorer health status, are less likely to use preventive services, are more likely to be hospitalised and are more likely to suffer poorer health outcomes for certain chronic conditions.[1]

The importance of health literacy has also been endorsed by a number of other agencies, which include the US Department of Health and Human Services, which identified improving health literacy as a specific objective in Healthy People 2010;[2] the European Commission, which cited the importance of health literacy in an influential 2002 health policy report;[3] and the World Health Organization, which included health literacy as a key component of global health promotion.[4]

In this paper, we will examine the epidemiology of health literacy, consider the range of skills associated with limited health literacy, review the research in both the adult and paediatric literature and propose strategies for clinicians who care for families with limited health-literacy skills.

EPIDEMIOLOGY

The Adult Literacy and Life Skills (ALL) survey was an international survey conducted in 2003 among seven countries or regions: Bermuda, Canada, Italy, Norway, Switzerland, the United States and the Mexican State of Nuevo Leon.[5] In the ALL survey, literacy and numeracy skills were described by five levels of proficiency. Experts considered level 3 literacy skills to be the minimum level necessary to meet the demands of an emerging knowledge society and information economy. In the ALL survey, prose literacy is defined as "the knowledge and skills needed to understand and use information from texts including editorials, new stories, brochures and instruction manuals".[5] Approximately 33% of adults in Norway, 40% of those in Canada and Bermuda, 50% of adults in Switzerland and the United States, 80% of those in Italy and almost 90% of adults in Nuevo Leon demonstrated less than level 3 prose literacy.[5] Additionally, countries varied significantly in the size of the gap in literacy skills between adults in the bottom 5% and those in the top 95%, with the gap being smaller in some countries (Norway and Switzerland) and larger in others (Italy).[5] As an international study of primarily developed nations, results of the ALL survey suggest that the effects of limited literacy probably affects individuals throughout the world.

In the ALL survey, lower literacy skills were associated with lower educational attainment, poorer economic outcomes, older age and poorer self-reported health status; these relationships have been supported by other studies.[1 5–8] Although literacy is associated with educational attainment, within each educational level individuals demonstrate a broad range of literacy skills.[5–7] Relying on educational attainment to predict adult literacy skills generally results in an overestimation of skill level.[7]

Although the United Kingdom did not participate in the ALL survey, it did participate in the International Adult Literacy Study (IALS), the predecessor of the ALL survey and the first international comparison of adult literacy skills.[9] Conducted between 1994 and 1998, the IALS assessed skills among 75 000 individuals in 22 nations.[9] In the IALS, 42% of Canadians, 47% of Americans and 52% of adults in the United Kingdom demonstrated below level 3 prose literacy skills.[9] In 2003, the British Department for

Education and Skills conducted the Skills for Life survey, a national literacy survey of English adults.[10] On the basis of this survey, 16% of English adults (5.2 million) had entry-level literacy skills, which were at or below the level expected of an 11 year old (ie, below a D–G grade in the General Certificate of Secondary Education qualification).[10] Given differences in the scoring and testing between the IALS and Skills for Life survey, it is difficult to draw direct comparisons between the outcomes of these two surveys. However, both of these studies suggest that the United Kingdom faces many of the same literacy issues as other developed nations.

In the United States, the 2003 National Assessment of Adult Literacy (NAAL), a nationally representative survey of adult English literacy skills, included an assessment of adult health literacy.[6] Limited health-literacy skills were associated with poverty, limited education, minority status, immigration and older age.[6] Results from the NAAL suggest that 36% of the US adult population have limited health-literacy skills.[6] This means that 77 million US adults may be unable to determine when to take a medication on the basis of directions associating medication administration with meals or may have difficulty identifying three substances that could interact with an over-the-counter medication based upon the drug label.[6]

In the NAAL, adults in the lowest health-literacy group were less likely to obtain health-related information from newspapers, books, magazines and the internet than were those with higher health literacy.[6] They were also less likely to obtain health-related information from family, friends, co-workers or healthcare providers than adults in higher health-literacy groups, but were more likely to obtain health-related information from the TV or radio.[6] These results highlight the importance of considering the accessibility of various forms of media when providing health education to individuals with lower literacy skills.

All countries participating in the ALL survey demonstrated a positive association between adult literacy skills and parental educational attainment, although the strength of this relationship varied between countries.[5] Studies of early literacy have found that maternal education is associated with a preschooler's engagement in home-based literacy activities such as reading together – an activity that has been associated with improved child language development and emergent early literacy skills.[11 12] Data from the 2003 NAAL suggest that parents with limited literacy skills were less likely to read to their young children five or more days a week and were less likely to have children between the ages of 3 and 5 who knew the alphabet.[13] Limited parental literacy may serve as a risk factor for the development of a child's emerging literacy skills.

The scope of health literacy

Although the majority of instruments used to measure health literacy focus on reading comprehension or numeracy skills, the concept of health literacy encompasses a wider array of abilities including reading, writing, listening and oral communication. Low literacy skills have been linked to poorer expressive communication, understanding and recall.[7] A recent study revealed that patients with low literacy levels were significantly less likely to ask questions, request additional services or seek new information during a medical encounter.[14] In the context of patient–provider communication, limitations in these skills may affect how a parent describes her child's symptoms, understands and remembers information or participates in medical decision-making.[7]

Poor literacy skills are often associated with a significant amount of shame and embarrassment. In one study, a third of patients with limited health-literacy skills denied any difficulties with reading.[15] Of those who acknowledged literacy issues, almost 40% endorsed feelings of shame – with the majority of this group having never told their spouse or children about their literacy challenges.[15]

Thus, in addition to the specific skills required by patients to make health-related decisions, one must consider the wider context in which patients acquire health-related information. An individual's ability to obtain, process and understand information is tied to the complexity of the information presented, the cultural overlay of health beliefs and the quality of patient–provider communication.[7] As such, many believe that the ''problem of health literacy'' is as much a problem of insufficient dedication of the staff within the healthcare system to the issue of reducing unnecessary complexity and communicating more effectively as it is a problem of limited literacy skills.[16]

With the increasing complexity of healthcare, the specialisation and technological advancements in medicine, as well as the growing reliance on self-management and home care in the treatment of chronic conditions, there are increasing literacy-based demands placed on patients and their families. In this environment, the role of health literacy probably assumes a greater contribution to health outcomes.

Health literacy and health outcomes in adults

Multiple studies have analysed the association between health literacy and health-related knowledge. Relationships between limited health-literacy skills and low levels of disease-specific knowledge have been demonstrated for a number of chronic conditions including asthma, diabetes, hypertension and congestive heart failure.[7 17]

Limited health literacy has been associated with decreased use of preventive health services such as immunisations and cancer screenings.[18] It has been associated with suboptimal disease-management skills as demonstrated by an increased likelihood of improper inhaler technique in adults with asthma and a decreased ability to appropriately identify medications in adults with coronary artery disease.[19–21] In addition, limited health-literacy skills have been consistently associated with worse health outcomes (ie, poorer physical functioning, poorer quality of life or late-stage disease detection) in conditions such as asthma and cancer,[22 23] increased rates of hospitalisation[24] and mortality.[25] Limited health-literacy skills, however, have yielded mixed results in diabetes, HIV and depression.[1 7 26–30] Similarly, studies of health literacy and medication adherence have also been inconsistent.[7 26 27 31]

Health literacy and child-health outcomes

Despite a significant body of health-literacy research in the adult medical literature, there have been few studies assessing the relationship between parental health literacy and child-health outcomes. A recent study demonstrated that limited parental health-literacy skills were associated with higher rates of ER visits, hospitalisations and severity of asthma symptoms for children with asthma.[32] Ross found that parental literacy, but not child literacy, was associated with HbA1c levels in children with type 1 diabetes.[33] Also, homeless mothers with limited health literacy were more likely to report barriers to giving their children the medication they needed than homeless mothers with adequate health literacy.[34]

By contrast, Sanders found no association between parental health literacy and a child's healthcare usage or healthcare costs among a population of urban children.[35] Similarly, in another study of children seen for acute outpatient care, there were no associations between parental health literacy and comprehension of a child's diagnosis or the ability to name and administer prescribed medications.[36]

In one study, parents with lower literacy levels attending a clinic staffed by residents actually reported higher-quality patient–provider relationships than did parents with higher literacy skills.[37] The authors posit that this finding does not represent a true difference in the quality of the patient–provider relationship, but instead suggests that individuals with lower literacy levels may have lower expectations and be less critical of their healthcare interactions.

Perhaps the effect of parental health literacy on child health will be more keenly observed in children with chronic conditions rather than in relatively healthy children for whom the degree of self-management is minimal. In addition, how parental health-literacy skills affect child-health outcomes may be influenced by the child's age and the child's own literacy skills. These and other questions regarding the role of parental health literacy in psediatric health outcomes continue to be an area of active research.

Measurement of health literacy

Most of the health-literacy instruments measure reading recognition, reading comprehension and/or numeracy skills. The Rapid Estimate of Adult Literacy in Medicine (REALM) and Test of Functional Health Literacy in Adults (TOFHLA) are two of the most commonly used health-literacy measures. These instruments, or a modified form of them, have been validated for use in adolescents.[38 39] The REALM and TOFHLA are primarily used for research purposes, and shortened versions of both tests are available. In addition, short screening tools, such as the ''Newest Vital Sign'' have been developed for use during office visits.[40]

Practitioners often have difficulty identifying individuals with limited literacy skills without the use of standardised instruments,[7 41] and developers of brief screening tools have promoted the idea of screening in the clinical setting.[40] We, however, question the usefulness of screening patients for limited health-literacy skills, noting that the only study examining the effect of a screening programme failed to show benefit for patients.[42] Currently, such testing could lead to stigma with no clear benefit. Instead, materials written in plain language, employing other clear communication techniques and confirming comprehension should be provided as part of medical care for all patients regardless of their literacy level.[43]

STRATEGIES FOR IMPROVING CARE FOR PATIENTS WITH LIMITED HEALTH LITERACY

Below we have listed some strategies that can be incorporated into a patient–provider interaction to enhance understanding and improve the quality of patient–provider communication.

Simplify the written word

The National Work Group on Literacy and Health recommends keeping written patient materials at the 5th grade level (generally 10–11 years of age) or lower and supplementing them with pictures or other forms of non-written communication.[44] Hundreds of studies have shown that the reading level of educational material for patient often exceeds the reading ability of its intended audience.[7] When assessing the readability of patient materials, clinicians might also consider documents such as appointment reminders, directions and patient intake forms as well as handouts for patient education.

Use strategies to improve comprehension and patient–provider communication

Clinicians can use a number of techniques to improve the quality of patient–provider communication. These strategies aim to improve communication by helping clinicians clarify their message and engage the parent in the decision-making process. Limiting the amount of information, using plain language, presenting recommendations as discrete action-oriented steps and assessing comprehension are some techniques suggested for improved communication.[45–48] In addition, engaging the patient in shared decision-making and exploring possible barriers to following recommendations may improve patient adherence and health outcomes. Box 1 provides descriptions of several patient–provider communication strategies.

Perform a ''literacy walk through''

Providers can evaluate the literacy demands placed on a parent bringing a child for a visit by performing a ''literacy walk through.''[54] Clinicians could consider the skills needed to navigate to the front desk, complete the sign-in process and participate in the visit. Attention should be devoted to signs, written information or directions that might be particularly challenging for families with limited health literacy.

Promote early child literacy and language skills

Reading to young children is an important part of developing early literacy and language skills. Programmes such as Reach-Out-And-Read, which encourage early literacy skills by providing books at well-child visits, can serve as an opportunity to model how books can be used to promote language development even if the parent's reading abilities are limited. Information regarding community-based adult-literacy programmes can also be provided to all families in a sensitive non-stigmatising manner. Benefits for children participating in Reach-Out-And-Read have been supported in the literature.[55]

DISCUSSION

Several experts in the field of health literacy have proposed mechanisms by which health-literacy skills may be linked to poorer health outcomes.[1 7 16 56] One conceptual model proposed by Paasche-Orlow and Wolf considers the contributions of both patient factors and system-based factors in influencing how limited health-literacy skills might be associated with the following three health-related processes: (1) access and utilisation of healthcare; (2) patient–provider communication; and (3) self-care management.[16] Within this framework, the mechanisms by which an individual's health-literacy skills interface with systemic demands can be used to consider areas of further research or possible interventions for families with limited health-literacy skills.

Consider the challenges a parent faces when navigating through a complex healthcare system. The ability to ''navigate'' includes both the tactical skills needed by an individual to successfully manoeuvre through a complicated healthcare centre filled with signs and placards as well as the ability to negotiate the complex web of regulations and bureaucracy found in many healthcare systems. Studies associating issues of

Box 1 Patient–provider communication strategies

Clinical skills (remove unnecessary complexities)
- Communicate using plain language
  - Use plain words and keep sentence structure simple.
- Avoid using jargon
  - Jargon can include both medical terminology (ie, anaemia) or words such as "diet", which has both lay and medical meanings. Both types of jargon can decrease comprehension.[49]
- Limit items discussed
  - Focus the discussion on the two or three most important ideas and reiterate these messages. Too much information or too many options may be overwhelming and sometimes result in decisions that are inconsistent with a patient's values.[50]
- Repeat important points

Be specific
- Provide clear, specific action-oriented steps
  - Information should answer the question, "What do I need to do?"
- Take the patient's perspective and consider what points remain unclear

Use multiple forms of communication
- Present the most important information through a variety of communication modalities
  - A combination of pictures, written and oral communication has been shown to improve understanding of health information in certain contexts.[51 52] Other forms of media such as video or interactive computer-based material may also be effective.
- Think like an educator!
  - Be creative and engage the learner.

Help patients to ask questions
- Provide an environment conducive to learning and asking questions
  - Parents with lower literacy may ask fewer questions and be less likely to describe possible barriers to recommendations. Lack of understanding can be shameful to the patient. Questions such as, "Do you understand?" may actually inhibit discussion. Consider instead, "I have asked you so many questions. What questions do you have for me?"
- "Ask Me 3"
  - Ask Me 3 is a campaign to promote health communication by having patients ask three questions in every healthcare encounter: (1) What is my main problem? (2) What do I need to do? and (3) Why is it important for me to do this? This campaign is sponsored by the Partnership for Clear Health Communication. Although research to assess the effectiveness of Ask Me 3 is ongoing, when we wrote this article, we were unable to find published studies. (www.askme3.org)

Learn to confirm comprehension
- Confirm comprehension with the "show me" or "teach back" method[53]
  - Ask the patient or parent to state or demonstrate the information presented. Using an iterative process, correct misunderstandings and have the patient or parent restate the information until comprehension is confirmed.
- Develop shared understanding
  - Explore the attitudes, beliefs and understanding of your patients and their families. Problem solve together about possible challenges or barriers to care.

access and healthcare use have been primarily conducted in the United States, whose complex multiple-payer system may pose a unique set of challenges when compared with the healthcare systems of other nations. Research conducted within other healthcare models may provide insight into how both individual health-literacy skills and systemic demands influence the association between health literacy and healthcare access and usage.

Patient–provider communication is another process through which health outcomes are probably affected. For example, little is known about how a parent's understanding of specific terminology affects the quality of a child's care. A study of parental comprehension of the word, "wheeze", suggested that parents who understood the meaning of the word "wheeze" were more likely to report "wheezing" in their children.[57] Providers typically fail to confirm that their patients understand new ideas or treatment plans, though doing so has been associated with better diabetes control.[53] Confirming comprehension can be done with a "show me" or "teach back" method in which patients demonstrate their understanding and clinicians give directed feedback to correct errors until mastery is established. For example, after teaching a family how to use an inhaler device, each person who will be participating in administering the medication can be asked to show how they will use the device (ie, "Show me how you plan to use this inhaler"). The clinician can observe the family's technique and provide focused feedback. This iterative process continues until the family demonstrates they have attained the goal. This teaching approach can improve shared understanding and decrease the chance of errors made through miscommunication.[53]

One of the most commonly proposed links between health literacy and poorer health outcomes has been through decreased self-care management skills, with significant interest in the role of medication adherence. Although limited health literacy has been associated with a decreased likelihood of appropriately identifying medications and doses,[20] studies of medication adherence have produced mixed results.[7 26 27 31] Perhaps other factors, such as family and social support mediate the relationship between health literacy and medication adherence. As researchers search for interventions to better support self-care management for various chronic conditions, understanding the role of these mediating variables may provide an important key to maximising self-care management for families with limited health literacy.

CONCLUSION

Research has demonstrated that limited health-literacy skills have important health-related implications for children and adults. Although the precise mechanism for the relationship between health literacy and health outcomes remains unknown, it is probably complex, involving individual and systemic factors. Without appropriate interventions, an individual's limited health-literacy skills may compromise his/her ability to engage fully in healthcare interactions. At a systemic level, the complexities of the medical system, the culture of medical care and the growing literacy-based demands placed on the patient may have substantial effects on the individual or family with limited health-literacy skills. The literature on health literacy should be a call to redirect significant resources not only toward patient education but also toward redesigning the system to promote and support self-care. Healthcare providers for children and the systems in which they work can have important roles in this process by assisting families with limited health-literacy skills and bridging the communication divide.

Acknowledgements: Dr. Hironaka receives support for her fellowship training from the National Research Service Award, Training Grant No. 2 T32 HP10014-13 and the Maternal Child Health Bureau Training Grant No. T77MC00016.

Funding: None.

Competing interests: None.

REFERENCES

1. **Nielsen-Bohlman LT,** Panzer A, Kindig D. *Health Literacy - A Prescription to End Confusion*. Washington DC: The National Academies Press, 2004.
2. **US Department of Health and Human Services**. Healthy People 2010: 2nd edn. *With understanding and improving health and objectives for improving health*. Washington DC: US Government Printing Office, 2000.
3. **European Commission**. *High-Level Group on innovation and provision of medicines in the European Union: Recommendations for action*. 2002. http://ec.europa.eu/health/ph_overview/Documents/key08_en.pdf (accessed 2 Feb 2007).
4. **World Health Organization**. *Health promotion: Report by the Secretariat*. Geneva: World Health Organization, 2000.
5. **Statistic Canada and Organization of Economic Co-operation and Development**. *Learning a living: First results of the adult literacy and life skills survey*. 2005. http://www.statcan.ca/english/freepub/89-603-XIE/2005001/pdf.htm (accessed 28 August 2007).
6. **Kutner M,** Greenberg E, Jin Y, *et al*. *The health literacy of America's adults - Results for the 2003 national assessment of adult literacy*. Washington, DC: US. Department of Education, 2006:i–60. http://nces.ed.gov/pubs2006/2006483.pdf (accessed 28 August 2007).
7. **Schwartzberg J,** Van Geest JB, Wang CC. Understanding health literacy: Implications for medicine and public health. Illinois: AMA Press, 2005.
8. **Paasche-Orlow MK,** Parker RM, Gazmararian JA, *et al*. The prevalence of limited health literacy. *J Gen Intern Med* 2005;**20**:175–84.
9. **Organization for Economic Co-operation and Development, Statistics Canada**. *Literacy in the information age - Final report of the international adult literacy survey*. 2000.http://www.oecdbookshop.org/oecd/display.asp?sf1 = identifiers&st1 = 812000051P1 (accessed 2 Feb 2008).
10. **Williams JC,** Clemens S, Oleinikova K, *et al*. *The Skills for Life survey - A national needs and impact survey of literacy, numeracy and ICT skills*. London, Department for Education and Skills, 2003. http://www.dfes.gov.uk/research/data/uploadfiles/RB490.pdf (accessed 2 Feb 2008).
11. **Raikes H,** Pan BA, Luze G, *et al*. Mother-child bookreading in low-income families: correlates and outcomes during the first three years of life. *Child Dev* 2006;**77**:924–53.
12. **Nord C,** Lennon J, Liu B, *et al*. Home literacy activities and signs of children's emerging literacy: 1993–1999. *Ed Stats Quart* 2000. http://nces.ed.gov/programs/quarterly/vol_2/2_1/q3-1.asp (accessed 2 Feb 2008).
13. **Kutner M,** Greenberg E, Jin Y, *et al*. *Literacy in everyday life: Results from 2003 National Assessment of Adult Literacy*. Washington DC: US Department of Education, 2007 http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid = 2007480 (accessed 2 Feb 2008).
14. **Katz MG,** Jacobson TA, Veledar E, *et al*. Patient literacy and question-asking behavior during the medical encounter: a mixed-methods analysis. *J Gen Intern Med* 2007;**22**:782–6.
15. **Parikh NS,** Parker RM, Nurss JR, *et al*. Shame and health literacy: the unspoken connection. *Patient Educ Couns* 1996;**27**:33–9.
16. **Paasche-Orlow MK,** Wolf MS. The causal pathways linking health literacy to health outcomes. *Am J Health Behav* 2007;**31**(Suppl 1):S19–26.
17. **Gazmararian JA,** Williams MV, Peel J, *et al*. Health literacy and knowledge of chronic disease. *Patient Educ Couns* 2003;**51**:267–75.
18. **Scott TL,** Gazmararian JA, Williams MV, *et al*. Health literacy and preventive health care use among Medicare enrollees in a managed care organization. *Med Care* 2002;**40**:395–404.
19. **Paasche-Orlow MK,** Riekert KA, Bilderback A, *et al*. Tailored education may reduce health literacy disparities in asthma self-management. *Am J Respir Crit Care Med* 2005;**172**:980–6.
20. **Kripalani S,** Henderson LE, Chiu EY, *et al*. Predictors of medication self-management skill in a low-literacy population. *J Gen Intern Med* 2006;**21**:852–6.
21. **Williams MV,** Baker DW, Honig EG, *et al*. Inadequate literacy is a barrier to asthma knowledge and self-care. *Chest* 1998;**114**:1008–15.
22. **Bennett CL,** Ferreira MR, Davis TC, *et al*. Relation between literacy, race, and stage of presentation among low-income patients with prostate cancer. *J Clin Oncol* 1998;**16**:3101–4.
23. **Mancuso CA,** Rincon M. Impact of health literacy on longitudinal asthma outcomes. *J Gen Intern Med* 2006;**21**:813–7.
24. **Baker DW,** Gazmararian JA, Williams MV, *et al*. Functional health literacy and the risk of hospital admission among Medicare managed care enrollees. *Am J Public Health* 2002;**92**:1278–83.
25. **Sudore RL,** Yaffe K, Satterfield S, *et al*. Limited literacy and mortality in the elderly: the health, aging, and body composition study. *J Gen Intern Med* 2006;**21**:806–12.
26. **Wolf MS,** Davis TC, Osborn CY, *et al*. Literacy, self-efficacy, and HIV medication adherence. *Patient Educ Couns* 2007;**65**:253–60.
27. **Paasche-Orlow MK,** Cheng DM, Palepu A, *et al*. Health literacy, antiretroviral adherence, and HIV-RNA suppression: a longitudinal perspective. *J Gen Intern Med* 2006;**21**:835–40.
28. **Schillinger D,** Grumbach K, Piette J, *et al*. Association of health literacy with diabetes outcomes. *JAMA* 2002;**288**:475–82.
29. **Morris NS,** MacLean CD, Littenberg B. Literacy and health outcomes: a cross-sectional study in 1002 adults with diabetes. *BMC Fam Pract* 2006;**7**:49.
30. **Gazmararian J,** Baker D, Parker R, *et al*. A multivariate analysis of factors associated with depression: evaluating the role of health literacy as a potential contributor. *Arch Intern Med* 2000;**160**:3307–14.
31. **Fang MC,** Machtinger EL, Wang F, *et al*. Health literacy and anticoagulation-related outcomes among patients taking warfarin. *J Gen Intern Med* 2006;**21**:841–6.
32. **DeWalt DA,** Dilling MH, Rosenthal MS, *et al*. Low parental literacy is associated with worse asthma care measures in children. *Ambul Pediatr* 2007;**7**:25–31.
33. **Ross LA,** Frier BM, Kelnar CJ, *et al*. Child and parental mental ability and glycaemic control in children with Type 1 diabetes. *Diabet Med* 2001;**18**:364–9.
34. **Sleath BL,** Jackson E, Thomas KC, *et al*. Literacy and perceived barriers to medication taking among homeless mothers and their children. *Am J Health Syst Pharm* 2006;**63**:346–51.
35. **Sanders LM,** Thompson VT, Wilkinson JD. Caregiver health literacy and the use of child health services. *Pediatrics* 2007;**119**:e86–92.
36. **Moon RY,** Cheng TL, Patel KM, *et al*. Parental literacy level and understanding of medical information. *Pediatrics* 1998;**102**:e25.
37. **Rosenthal MS,** Socolar RR, DeWalt DA, *et al*. Parents with low literacy report higher quality of parent-provider relationships in a residency clinic. *Ambul Pediatr* 2007;**7**:51–5.
38. **Chisolm DJ,** Buchanan L. Measuring adolescent functional health literacy: a pilot validation of the Test of Functional Health Literacy in Adults. *J Adolesc Health* 2007;**41**:312–4.
39. **Davis TC,** Wolf MS, Arnold CL, *et al*. Development and validation of the Rapid Estimate of Adolescent Literacy in Medicine (REALM-Teen): a tool to screen adolescents for below-grade reading in health care settings. *Pediatrics* 2006;**118**:e1707–14.
40. **Weiss BD,** Mays MZ, Martz W, *et al*. Quick assessment of literacy in primary care: the newest vital sign. *Ann Fam Med* 2005;**3**:514–22.
41. **Wittich AR,** Mangan J, Grad R, *et al*. Pediatric asthma: caregiver health literacy and the clinician's perception. *J Asthma* 2007;**44**(1):51–5.
42. **Seligman HK,** Wang FF, Palacios JL, *et al*. Physician notification of their diabetes patients' limited health literacy. A randomized, controlled trial. *J Gen Intern Med* 2005;**20**:1001–7.
43. **Paasche-Orlow MK,** Schillinger D, Greene SM, *et al*. How health care systems can begin to address the challenge of limited literacy. *J Gen Intern Med* 2006;**21**:884–7.
44. **Communicating with patients who have limited literacy skills**. Report of the National Work Group on Literacy and Health. *J Fam Pract* 1998;**46**:168–76.
45. **Paasche-Orlow MK,** Parker RM. Improving the effectiveness of patient education: A focus on limited health literacy. In: King TE, Wheeler M, Fernandez A, *et al.*, eds. *Medical Management of Vulnerable and Underserved Patients: Principles, Practice, and Populations*. New York: McGraw-Hill, 2007.
46. **Doak CC,** Doak LG, Friedell GH, *et al*. Improving comprehension for cancer patients with low literacy skills: strategies for clinicians. *CA Cancer J Clin* 1998;**48**:151–62.
47. **Mayeaux EJ Jr,** Murphy PW, Arnold C, *et al*. Improving patient education for patients with low literacy skills. *Am Fam Physician* 1996;**53**:205–11.
48. **Williams MV,** Davis T, Parker RM, *et al*. The role of health literacy in patient-physician communication. *Fam Med* 2002;**34**:383–9.
49. **Castro CM,** Wilson C, Wang F, *et al*. Babel Babble: Physicians' use of unclarified medical jargon with patients. *Am J Health Behav* 2007;**31**(Suppl.1):585–595.
50. **Ubel PA.** Is information always a good thing? Helping patients make ''good'' decisions. *Med Care* 2002;**40**(9 Suppl):V39–44.
51. **Houts PS,** Doak CC, Doak LG, *et al*. The role of pictures in improving health communication: a review of research on attention, comprehension, recall, and adherence. *Patient Educ Couns* 2006;**61**:173–90.
52. **Katz MG,** Kripalani S, Weiss BD. Use of pictorial aids in medication instructions: a review of the literature. *Am J Health Syst Pharm* 2006;**63**:2391–7.
53. **Schillinger D,** Piette J, Grumbach K, *et al*. Closing the loop: physician communication with diabetic patients who have low health literacy. *Arch Intern Med* 2003;**163**:83–90.
54. **Hixon AL.** Functional health literacy: improving health outcomes. *Am Fam Physician* 2004;**69**:2077–8.
55. **Willis E,** Kabler-Babbitt C, Zuckerman B. Early literacy interventions: reach out and read. *Pediatr Clin North Am* 2007;**54**:625–42, viii.
56. **Baker DW.** The meaning and the measure of health literacy. *J Gen Intern Med* 2006;**21**:878–83.
57. **Michel G,** Silverman M, Strippoli MP, *et al*. Parental understanding of wheeze and its impact on asthma prevalence estimates. *Eur Respir J* 2006;**28**:1124–30.

Thrombosis Research (2008) **123**, S46–S49

THROMBOSIS RESEARCH

www.elsevier.com/locate/thromres

REGULAR ARTICLE

Oral anticoagulant therapy for patients with atrial fibrillation - an update

Elaine M. Hylek

Boston University School of Medicine, Department of Medicine, Boston, MA

Available online 21 September 2008

Epidemiology of Atrial Fibrillation

It is projected that 7.5 million individuals will have atrial fibrillation (AF) in the United States by the year 2020 [1]. The reason for this increasing prevalence is multifactorial and due in part to the increased prevalence of risk factors for AF like hypertension, heart failure, diabetes, obesity, and older age. Among individuals age 80 years and greater, the prevalence of AF is approximately 10 percent [2]. Atrial fibrillation confers a five-fold increased risk of stroke and 15 percent of all strokes occur in people with AF. The attributable risk of stroke with AF differs by age, 1.5 percent for those 50 to 59 years versus 23.5 percent for those 80 to 89 years. Untreated, the average rate of stroke per year is 5 percent [3]. The global impact of stroke is enormous. It is the third most common cause of death in developed countries and the leading cause of serious long term disability. The disability adjusted life years are projected to reach 61 million years per 1,000 population by the year 2020 which is significantly increased from 38 million in 1990 [4]. The 30-day mortality of AF-related stroke has been shown in multiple studies to approximate 23-25 percent [5].

Stroke Prevention in Atrial Fibrillation

Oral vitamin K antagonists like warfarin are extremely effective in reducing the risk of stroke among patients with AF. The early AF trials consistently demonstrated benefit with an overall 68% risk reduction in stroke [6]. Despite this dramatic efficacy, warfarin remains underused in clinical practice. A study of 21 academic, 13 community, and 4 Veterans Administration hospitals in the U.S. revealed that 53% of individuals with AF and without obvious contraindications were prescribed warfarin at discharge. Age greater than 80 years and perceived bleeding risk were the most often cited reasons for not prescribing warfarin [7]. The gap is even greater among economically disadvantaged groups [8]. For example, among Ohio Medicaid patients, only 11.9 percent of patients without documented contraindications filled a prescription for warfarin based on a retrospective analysis of 1998-2000 claims data. Barriers to compliance either with laboratory monitoring or medication adherence accounted for approximately 30 percent [9].

Intracranial hemorrhage is the most feared complication of anticoagulant therapy. Mortality related to intracerebral hemorrhage approximates 50 percent and is related to hematoma volume and hematoma expansion [10]. Risk factors for intracranial hemorrhage include anticoagulation intensity, age, prior stroke, hypertension, concomitant antiplatelet therapy, and arterial vasculopathy, specifically amyloid

E-mail address: ehylek@bu.edu.

0049-3848/$ - see front matter
doi:10.1016/j.thromres.2008.08.016

angiopathy and leukoaraiosis. The paradox facing clinicians and patients is that many of the risk factors for intracranial hemorrhage are also risk factors for ischemic stroke. Age, hypertension, and prior stroke all increase the risk for both hemorrhagic and non-hemorrhagic stroke. Although leukoaraiosis and amyloid angiopathy are associated with increased risk of intracerebral hemorrhage, their utility for risk stratification remains unclear.

When assessing risk versus benefit, current evidence from randomized controlled trials and observational studies weighs in favor of anticoagulant therapy. The rate of intracranial hemorrhage is approximately 0.5-0.6 percent. The rate of major extracranial hemorrhage (defined as fatal, necessitating transfusion of two or more units of packed cells, or occurring in a critical site) is approximately 2 percent [11,12]. Untreated, the rate of stroke among individuals with AF and multiple risk factors exceeds 10 percent.

However, there are important caveats to the published rates of major hemorrhage. Overall, few patients greater than 80 years of age have been enrolled in randomized trials and cohort studies. In addition, recent trials and most observational cohorts have largely reported the outcomes of prevalent users of warfarin. Because most major bleeding occurs early in the course of anticoagulation, enrollment of predominantly long-time users of warfarin will result in lower estimates of hemorrhage. This survival bias was illustrated by an inception cohort study of patients with AF aged 65 years and greater who were newly starting warfarin therapy. The rate of major hemorrhage in the first year was 7 percent and discontinuation of warfarin was highest in the first 90 days. Perhaps not surprisingly, cessation of warfarin was disproportionately high among patients at the highest risk of stroke [13].

Strategies to Optimize the Benefit of Oral Anticoagulant Therapy

Anticoagulation Control

Vitamin K antagonists like warfarin are remarkably efficacious in preventing stroke. However, their narrow therapeutic window, variable dose response, and interference by diet and other medications mandate frequent monitoring of the International Normalized Ratio (INR). As an example, amiodarone necessitates reduction of maintenance warfarin dose by approximately one-third in many patients due to its interference with the metabolism of both enantiomers of warfarin. Protracted use of high doses of acetaminophen has been shown to interfere with the enzymes of the vitamin K cycle. Certain disease states, most notably decompensated heart failure and active cancer, induce significant variability in the INR and these patients warrant frequent monitoring [14]. Warfarin dose requirements decrease by approximately 50 percent with advancing age which highlights the importance of initiating warfarin with smaller doses in elderly patients [15]. Older patients are also slower to normalize an elevated INR which may in part underlie the increased bleeding seen in this age group. In a study of 633 patients presenting with an INR greater than 6.0, each decade of age conferred an 18% increased risk of having an INR greater than 4.0 on repeat testing after withholding 2 consecutive doses of warfarin. Decompensated heart failure, chemotherapy, and a lower warfarin maintenance dose all predicted a slower return to the therapeutic range [16].

Recent interest has grown in the potential role of genetic polymorphisms in predicting warfarin dose, but definitive guidance awaits randomized trials. Because so many factors affect warfarin dose, patients will continue to need frequent monitoring in the first 2 weeks of initiation regardless of genetic profile. In addition, dietary patterns and medications frequently change following hospital discharge. Understanding the pharmacokinetics of warfarin, triggers for aberrant control, influence of age, and the effects of illness on INR and warfarin dose will help to improve anticoagulation control in the interim.

Lower INR Target Intensity

There are few data to support use of a lower INR target intensity for patients with AF. Sub-therapeutic INR levels of less than 2.0 have been associated with more thromboembolic events and more severe strokes [17]. Although an INR target of 2.0-2.5 seems like a rational compromise, this degree of precision may be difficult to attain in clinical practice. Among patients with recurrent hemorrhage, the balance of risk versus benefit of a lower target range has never been studied. Patients and their caregivers need to be apprised of the potential trade-offs inherent to such a strategy.

Concurrent Antiplatelet Therapy

There is a growing appreciation for the hazards of combination therapy particularly among older patients. A recent meta-analysis highlighted the heightened bleeding risk [18]. The investigators analyzed data from 10 clinical trials that compared oral anticoagulant therapy alone versus oral

anticoagulant therapy plus aspirin among 4,180 patients with AF, mechanical heart valve or coronary artery disease. Combination therapy was associated with a lower incidence of arterial thromboembolism (odds ratio = 0.66), but the benefits were limited to patients with mechanical valves (odds ratio = 0.27). Combination therapy did not benefit patients with AF (odds ratio = 0.99) or coronary artery disease (odds ratio = 0.69) nor did it influence all cause mortality. Oral anticoagulant therapy plus aspirin did increase the risk of major bleeding (odds ratio = 1.43). Aspirin has also been shown to increase the risk of intracerebral hemorrhage. The role of aspirin among patients taking warfarin for AF who have stable coronary artery disease is currently being reassessed. Strategies are also evolving for patients with coronary stents and AF who require warfarin and dual antiplatelet therapy with aspirin and clopidogrel.

Blood Pressure Control

An often overlooked yet critical issue for patients on oral anticoagulant therapy is blood pressure control. Blood pressure less than 130/80 is associated with a decrease in both ischemic and hemorrhagic stroke [19,20]. Uncontrolled hypertension is a contraindication to warfarin therapy.

Risk Stratification for Stroke

Several risk stratification schemes have been published. Perhaps the most widely used is the $CHADS_2$ score which assigns points to stroke risk factors (C=Congestive heart failure, H=Hypertension, A=Age $\geq$ 75 years, D=Diabetes mellitus, S=prior Stroke/systemic embolus/transient ischemic attack). Each risk factor is assigned one point except prior stroke which is 2 points [21]. As an example, the $CHADS_2$ score for an 80 year old woman with a history of transient ischemic attack and hypertension=4. Current guidelines recommend warfarin for a $CHADS_2$ score of 2 or greater and either warfarin or aspirin for a $CHADS_2$ score=1 depending on patient preference and ease of INR monitoring [22]. The $CHADS_2$ scheme was derived from a Medicare population, as opposed to patients enrolled in randomized trials thereby increasing the external validity of the tool. However, the ability of these schemes to predict stroke outside of the derivation population remains suboptimal. Definitions for key covariates have differed and only the Framingham model incorporates blood pressure measurement instead of simply the presence or absence of hypertension as a diagnosis. Better models to predict stroke and hemorrhage are needed.

New Antithrombotic Drugs

There is a pressing need for antithrombotic drugs with a wider therapeutic index, decreased potential for interactions with diet and other medications, and minimal need for monitoring. Ximelagatran, an oral direct thrombin inhibitor, was the first of these agents. It was withdrawn due to hepatic toxicity. Multiple other promising new drugs are currently undergoing evaluation in randomized controlled trials and Phase II testing. The size of the trial populations necessary to evaluate the efficacy of these new drugs versus warfarin is daunting. The optimal double-blind design is challenging logistically for many of the countries and sites participating in these trials. In addition, because the risk of hemorrhage and stroke are greatest in the early period, studying adequate numbers of patients newly taking these drugs will be important.

Control of AF

Extensive research into the basic mechanisms underlying AF and propagation of AF are ongoing. The role of inflammation, vascular stiffness, oxidant stress, nitric oxide, and the renin-aldosterone angiotensin system (RAAS) have yet to be elucidated. A recent study from the Framingham cohort linked incidence of AF with pulse pressure (surrogate for vascular stiffness). The investigators demonstrated earlier progression to AF among those patients in the highest quartile of pulse pressure [23].

In a meta-analysis of 11 studies involving either angiotensin converting enzyme inhibitors or angiotensin receptor blockers, Healy, et al., demonstrated a decreased incidence of AF among patients randomized to these treatment arms. Overall the risk of AF was reduced by 28 percent and by 44% among patients with left ventricular dysfunction or left ventricular hypertrophy [24]. This finding awaits validation by randomized controlled trials. Regarding non-pharmacological strategies, the role of pulmonary vein isolation in the management of AF is evolving as is identification of the optimal candidates for this procedure.

Conflict of interest statement

I have received research grants from AstraZeneca, Bristol-Myers Squibb, and Pfizer, and currently

participate in an advisory capacity for Bristol-Myers Squibb.

References

[1] Miyasaka Y, Barnes ME, Gersh BJ, Cha SS, Bailey KR, Abhayaratna WP, Seward JB, Tsang TSM. Secular Trends in Incidence of Atrial Fibrillation in Olmstead County, Minnesota, 1980-2000, and Implications on the Projections for Future Prevalence. *Circulation* 2006;**114:**119–25.

[2] Wolf PA, Abbott RD, Kannel WB. Atrial fibrillation as an independent risk factor for stroke: the Framingham Study. *Stroke* 1991;**22(8):**983–8.

[3] Benjamin EJ, Wolf PA, D'Agostino RB, Silbershatz H, Kannel WB, Levy D. Impact of Atrial Fibrillation on the Risk of Death: The Framingham Heart Study. *Circulation* 1998;**98(10):**946–52.

[4] Mackay J, Mensah GA. The Atlas of Heart Disease and Stroke. Geneva, Switzerland: World Health Organization; 2004.

[5] American Heart Association. Heart Disease and Stroke Statistics–2006 Update. *Circulation* 2006;**113(6):**e85–e151.

[6] Atrial Fibrillation Investigators. 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):**1449–57.

[7] Waldo AL, Becker RC, Tapson VF, Colgan KJ. Hospitalized Patients With Atrial Fibrillation and a High Risk of Stroke Are Not Being Provided With Adequate Anticoagulation. *J Am Coll Cardiol* 2005;**46(9):**1729–36.

[8] Birman-Deych ERM, Nilasena DS, Gage BF. Use and effectiveness of warfarin in Medicare beneficiaries with atrial fibrillation. *Stroke* 2006;**37(4):**1070–4.

[9] Johnston JA, Cluxton Jr RJ, Heaton PC, Guo JJ, Moomaw CJ, Eckman MH. Predictors of Warfarin Use Among Ohio Medicaid Patients With New-Onset Nonvalvular Atrial Fibrillation. *Arch Intern Med* 2003;**163(14):**1705–10.

[10] Davis SM, Broderick J, Hennerici M, Brun NC, Diringer MN, Mayer SA, Begtrup K, Steiner T, for the Recombinant Activated Factor VIIIHTI. Hematoma growth is a determinant of mortality and poor outcome after intracerebral hemorrhage. *Neurology* 2006;**66:**1175–81.

[11] Go AS, Hylek EM, Chang Y, Phillips KA, Henault LE, Capra AM, Jensvold NG, Selby JV, Singer DE. Anticoagulation Therapy for Stroke Prevention in Atrial Fibrillation: How Well Do Randomized Trials Translate Into Clinical Practice? *JAMA* 2003;**290(20):**2685–92.

[12] The Atrial Fibrillation Follow-up Investigation of Rhythm Management I. A Comparison of Rate Control and Rhythm Control in Patients with Atrial Fibrillation. *N Engl J Med* 2002;**347(23):**1825–33.

[13] Hylek EM, Evans-Molina C, Shea C, Henault LE, Regan S. Major Hemorrhage and Tolerability of Warfarin in the First Year of Therapy Among Elderly Patients With Atrial Fibrillation. *Circulation* 2007;**115(21):**2689–96.

[14] Ansell J, Hirsh J, Poller L, Bussey H, Jacobson A, Hylek EM. The pharmacology and management of the vitamin K antagonists: The seventh ACCP conference on antithrombotic and thrombolytic therapy. *Chest* 2004;**126:**204S–33S.

[15] 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):**2049–56.

[16] Hylek EM, Regan S, Go AS, Hughes RA, Singer DE, Skates SJ. Clinical Predictors of Prolonged Delay in Return of the International Normalized Ratio to within the Therapeutic Range after Excessive Anticoagulation with Warfarin. *Ann Intern Med* 2001;**135(6):**393–400.

[17] Hylek EM, Go AS, Chang Y, Jensvold NG, Capra AM, Phillips KA, Henault LE, Selby JV, Singer DE. Effect of intensity of oral anticoagulation on stroke severity and mortality in atrial fibrillation. *N Engl J Med* 2003;**349(11):**1019–26.

[18] 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):**117–24.

[19] Lip GY, Frison L, Grind M. Effect of hypertension on anticoagulated patients with atrial fibrillation. *Eur Heart J* 2007;**28(6):**752–9.

[20] Arima H, Hart RG, Colman S, Chalmers J, Anderon C, Rodgers A, Woodward M, MacMahon S, Neal B, for the 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:**2164–9.

[21] 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):**2864–70.

[22] Fuster V, Ryden LE, Cannom DS, et al. ACC/AHA/ESC 2006 Guidelines for the Management of Patients With Atrial Fibrillation–Executive Summary. *Circulation* 2006;**114(7):** 700–52.

[23] Mitchell GF, Vasan RS, Keyes MJ, Parise H, Wang TJ, Larson MG, D'Agostino Sr RB, Kannel WB, Levy D, Benjamin EJ. Pulse Pressure and Risk of New-Onset Atrial Fibrillation. *JAMA* 2007;**297(7):**709–15.

[24] Healey JS, Baranchuk A, Crystal E, Morillo CA, Garfinkle M, Yusuf S, Connolly SJ. Prevention of Atrial Fibrillation With Angiotensin-Converting Enzyme Inhibitors and Angiotensin Receptor Blockers: A Meta-Analysis. *Journal of the American College of Cardiology* 2005;**45(11):**1832–9.

Contra: 'Warfarin should be the drug of choice for thromboprophylaxis in elderly patients with atrial fibrillation'

Caveats regarding use of oral anticoagulant therapy among elderly patients with atrial fibrillation

Elaine M. Hylek

Research Unit-Section of General Internal Medicine and Clinical Epidemiology Unit, Boston University School of Medicine, Boston, Massachussetts, USA

Although the balance of risk versus benefit weighs in favor of anticoagulation therapy for most elderly individuals with atrial fibrillation (AF), the potential for harm is considerably greater in the older age group. Higher rates of major haemorrhage have repeatedly been shown among older individuals in both randomized trials and observational studies (1–4). In the AFFIRM trial (Atrial Fibrillation Follow-up Investigation of Rhythm Management), the risk of major bleeding increased by approximately 5% per year of age (3). Most recently, elderly patients and those with renal impairment were found to be at greatest risk for bleeding complications in the trial with idraparinux, a once weekly injectable factor Xa inhibitor (5).

To mitigate the risk of bleeding in the elderly population, it is critically important to understand the factors that underlie this risk. Age alone is not a contraindication to anticoagulant therapy. The spectrum of aging is broad and has been categorized into three distinct groups: the young-old (age 65–74 years), old-old (age 75 to 84), and oldest-old (age 85 and over). Functional status varies considerably across and within these different age strata as does frailty and the risk of haemorrhage. In randomized trials, eligibility criteria and rigorous follow-up protocols tend to bias enrollment to the highest functioning and overall healthier elderly patients, suggesting that biological age probably means more than chronological age. In addition, most trials in AF have excluded patients with either an indication for antiplatelet therapy or a perceived contraindication to anticoagulation. There is little consensus on the factors that constitute an absolute contraindication, and designation of a relative contraindication to anticoagulant therapy is largely subjective. The systematic review as part of the UK National Institute for Health and Clinical Excellence (NICE) management guidelines for AF identified the following patient characteristics as risk factors for anticoagulation-related bleeding complications: advanced age, uncontrolled hypertension, history of myocardial infarction or ischaemic heart disease, cerebrovascular disease, anaemia or a history of bleeding, and the concomitant use of other drugs such as antiplatelet drugs (6). Importantly, some of the risk factors for anticoagulation-related bleeding were also factors indicating the use of anticoagulants in AF patients.

In stroke prevention trials, the nearly universal clinical trial restrictions on aspirin (or clopidogrel) use facilitates enrollment of patients with a lower atherosclerotic burden which further lowers the risk. To more fully interpret the experience of elderly trial participants, knowledge of the denominator and number of individuals screened in these respective age categories is important. All of these study design features contribute to the disparate rates of haemorrhage reported from trials and real-world observational cohorts. The initial selection bias related to warfarin candidacy in trials and in clinical practice is a culmination of patient and physician judgment and preference. The impact of this bias on risk is difficult to quantify and thwarts attempts to extrapolate rates of haemorrhage across different patient populations.

Aging is accompanied by pathological changes that increase the propensity for haemorrhage. The incidence of lower gastrointestinal haemrrhage increases 200-fold from age 30 to age 90 years (7). Diverticulosis, malignancy, angiodysplasias, and ischemic colitis are all more common in the elderly population. Recurrent bleeding especially from small bowel vascular ectasias makes definitive treatment challenging and often precipitates cessation of anticoagulant therapy. The risk of upper gastrointestinal bleeding or perforated peptic ulcer with antiplatelet therapy increases exponentially with age (8). These detrimental effects have stimulated interest in nitric oxide (NO)-donating aspirin to evaluate the touted protective effect of NO on the gastrointestinal mucosa (9). Pathological changes in cerebral vessels, specifically leukoaraiosis and amyloid angiopathy, in-

Correspondence to:
Elaine M. Hylek, MD, MPH
Research Unit-Section of General Internal Medicine and Clinical Epidemiology Unit
Boston University School of Medicine
801 Massachusetts Avenue, 2nd floor
Boston, MA 02118, USA
Tel.: +1 617 414 3743, Fax: +1 617 414 4676
E-mail: ehylek@bu.edu

Received June 1, 2008
Accepted June 1, 2008

Prepublished online June 11, 2008
doi:10.1160/TH08-06-0343

Thromb Haemost 2008; 100: 16–17

crease with age and predispose to intracerebral haemorrhage (10, 11). The increased risk of subdural bleeding, both spontaneous and associated with minor trauma, reflects age-related changes in bridging dural vessels. The increased risk of falls among elderly individuals is multifactorial and, in part, due to sensory impairment, sarcopenia, autonomic dysfunction, and orthostatic hypotension. The latter is often iatrogenic and secondary to the detrimental effects of polypharmacy, for example, alpha-blockers used in conjunction with diuretics, nodal agents, and/or vasodilators. In one study, the rate of intracranial haemorrhage among individuals prone to fall was 2.8% (95% confidence interval, 1.9–4.1) (12).

Several studies have now highlighted one of the key clinical dilemmas in stroke prevention among patients with atrial fibrillation, i.e. patients at highest risk of stroke ($CHADS_2$ score ≥3) also experience most of the major bleeding complications on antithrombotic therapy (4, 5, 13, 14). This trend was also evident among the patients randomized to idraparinux in the Amadeus trial and clopidogrel plus aspirin in the ACTIVE-W trial, affirming that the increased bleeding complications were not attributable to greater international normalised ratio (INR) variability among the higher risk warfarin-treated patients. Death and disability associated with anticoagulant therapy are almost exclusively due to intracranial haemorrhage (15). However, even in the older age group, the risk of intracranial bleeding is dwarfed by the risk of ischemic stroke. Cardioembolic strokes are associated with a 30-day mortality of 24% (16, 17). Although extracranial haemorrhage is not associated with the same morbidity and mortality of intracranial haemorrhage, it often precipitates cessation of warfarin therapy.

The debate has therefore shifted from whether or not elderly patients are at higher risk of haemorrhage to what are the optimal strategies to minimize its occurrence. The hazards of antiplatelet therapy combined with oral anticoagulants have been increasingly realized (18). Blood pressure control is critical as levels of 130/80 or less have been shown to decrease the risk of both ischemic and haemorrhagic stroke (19, 20). In addition, elderly patients require considerably lower doses of warfarin to attain the target range and are slower to recover from an episode of excessive anticoagulation (21). Because the initial phase of warfarin therapy is the most risk-prone period, frequent monitoring is therefore essential. Whether or not newer anticoagulant drugs with considerably shorter half-lives, wider therapeutic margins, and less dietary and drug interactions will be safer in elderly patients with atrial fibrillation awaits results of ongoing clinical trials.

References

1. van der Meer FJ, Rosendaal FR, Vandenbroucke JP, et al. Bleeding complications in oral anticoagulant therapy. An analysis of risk factors. Arch Intern Med 1993; 153: 1557–1562.
2. The Stroke Prevention in Atrial Fibrillation Investigators. Bleeding during antithrombotic therapy in patients with atrial fibrillation. Arch Intern Med 1996; 156: 409–416.
3. DiMarco JP, Flaker G, Waldo AL, et al., and the AFFIRM Investigators. Factors affecting bleeding risk during anticoagulant therapy in patients with atrial fibrillation. Am Heart J 2005; 149: 650–656.
4. Hylek EM, Evans-Molina C, Shea C, et al. Major hemorrhage and tolerability of warfarin in the first year of therapy among elderly patients with atrial fibrillation. Circulation 2007; 115: 268–2696.
5. The Amadeus Investigators. Comparison of idraparinux with vitamin K antagonists for prevention of thromboembolism in patients with atrial fibrillation: a randomized, open-label, non-inferiority trial. Lancet 2008; 371: 315–321.
6. 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: 295–304.
7. Longstreth GF. Epidemiology and outcome of patients hospitalized with acute lower gastrointestinal hemorrhage: A population-based study. Am J Gastroenterol 1997; 92: 419–424.
8. Hernandez-Diaz S, Garcia Rodriguez LA. Cardioprotective aspirin users and their excess risk of upper gastrointestinal complications. BMC Med 2006; 4: 22–28.
9. Bolla M, Momi S, Gresele P, et al. Nitric oxide-donating aspirin (NCX 4016): an overview of its pharmacological properties and clinical perspectives. Eur J Clin Pharmacol 2005; 26: 1–10.
10. Rosand J, Hylek EM, O'Donnell HC, et al. Warfarin-associated hemorrhage and cerebral amyloid angiopathy: a genetic and pathologic study. Neurology 2000; 55: 947–951.
11. Smith EE, Rosand J, Knudsen KA, et al. Leukoaraiosis is associated with warfarin-related hemorrhage following ischemic stroke. Neurology 2002; 59: 193–197.
12. Gage BF, Birman-Deych E, Kerzner R, et al. Incidence of intracranial hemorrhage in patients with atrial fibrillation who are prone to fall. Am J Med 2005; 118: 612–617.
13. Poli D, Antonucci E, Marcucci R, et al. Risk of bleeding in very old atrial fibrillation patients on warfarin: Relationship with ageing and $CHADS_2$ score. Thromb Res 2007; 121: 347–352.
14. 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. Stroke 2008; 39: 1–5.
15. Fang MC, Go AS, Chang Y, et al. Death and disability from warfarin-associated intracranial and extracranial hemorrhages. Am J Med 2007; 120: 700–705.
16. Lin H, Wolf PA, Kelly-Hayes M, et al. Stroke Severity in Atrial Fibrillation: The Framingham Study. Stroke 1996; 27: 1760–1764.
17. 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: 1019–1026.
18. Dentali F, Douketis JD, Lim W, et al. 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: 117–124.
19. Hart RG, Tonarelli SB, Pearce LA. Avoiding central nervous system bleeding during antithrombotic therapy: recent data and ideas. Stroke 2005; 36: 1588–1593.
20. Lip GY, Frison L, Grind M. Effect of hypertension on anticoagulated patients with atrial fibrillation. Eur Heart J 2007; 28: 752–759.
21. Garcia D, Regan S, Crowther M, et al. Warfarin maintenance dosing patterns in clinical practice: Implications for safer anticoagulation in the elderly population. Chest 2005; 127: 2049–2056.

Ultimately, it is likely that the field of aging and dementia research will move in the direction of using a combination of imaging measures, biomarkers, and subtle clinical features to determine who might be at the greatest risk for progressing to dementia. For example, one might consider a person's family history, genetic status, a blood biomarker, and medical comorbidities, and combine these factors with a memory measure to stratify risk. More expensive and invasive investigations, such as a magnetic resonance imaging scan, cerebrospinal fluid analysis, and amyloid imaging measures, might be entertained for those at a greater risk according to this profile. The data that Amieva and colleagues[7] provided give insight into the type of clinical measures to be considered.

Finally, the context in which biomarker profiles including cognitive tests will be used in the early detection and prevention of AD is an important concern. The incorrect identification of individuals as being at high risk for AD may lead to undue alarm and concern. Therefore, before biomarker profiles are used in the general population, high specificity should be demonstrated in multiple populations. The initial application of biomarkers and cognitive tests for early detection is likely to be for individuals coming to the clinic with concerns about their memory and for alerting the clinician of the need to carefully follow patients who have profiles indicating high risk for AD. When safe and effective disease-modifying therapies have been developed, reliable early detection in the general population will become an essential tool in the prevention of this illness.

James A. Mortimer, PhD[1] and
Ronald C. Petersen, MD, PhD[2]

[1]Department of Epidemiology and Biostatistics College of Public Health, University of South Florida
Tampa, FL

[2]Department of Neurology
Alzheimer's Disease Research Center, Mayo Clinic College of Medicine
Rochester, MN

Potential conflict of interest: Nothing to report.

References

1. Ferri CP, Prince M, Brayne C, et al. Global prevalence of dementia: a Delphi consensus study. Lancet 2005;366: 2112–2117.
2. Sloane PD, Zimmerman S, Suchindran C, et al. The public health impact of Alzheimer's disease, 2000-2050: potential implication of treatment advances. Annu Rev Public Health 2002; 23:213–231.
3. Brookmeyer R, Gray S, Kawas C. Projections of Alzheimer's disease in the United States and the public health impact of delaying disease onset. Am J Public Health 1998;88: 1337–1342.
4. DeKosky ST, Marek K. Looking backward to move forward: early detection of neurodegenerative disorders. Science 2003; 302:830–834.
5. Cummings, JL, Doody R, Clark C. Disease-modifying therapies for Alzheimer's disease: challenges to early intervention. Neurology 2007;69:1622–1634.
6. Shaw LM, Korecka M, Clark CM, et al. Biomarkers of neurodegeneration for diagnosis and monitoring therapeutics. Nat Rev Drug Discov 2007;6:295–303.
7. Amieva H, Goff ML, Millet X, et al. Prodromal Alzheimer's disease: Successive emergence of the clinical symptoms. Ann Neurol 2008;64:492–498.
8. Elias MF, Beiser A, Wolf PA, et al. The preclinical phase of Alzheimer's disease: a 22-year prospective study of the Framingham cohort. Arch Neurol 2000;57:800–813.
9. Kawas C, Corrada MM, Brookmeyer R, et al. Visual memory predicts Alzheimer's disease more than a decade before diagnosis. Neurology 2003;60:1089–1093.
10. Welsh K, Butters N, Hughes J, et al. Detection of abnormal memory decline in mild cases of Alzheimer's disease using CERAD neuropsychological measures. Arch Neurol 1991;48: 278–281.
11. Petersen RC, Smith G, Kokmen E, et al. Memory function in normal aging. Neurology 1992;42:396–401.
12. Chapman LJ, Chapman JP. The measurement of differential deficit. J Psychiat Res 1978;14:303–311.
13. Backman L, Jones S, Berger A-K, et al. Cognitive impairment in preclinical Alzheimer's disease: a meta-analysis. Neuropsychology 2005;19:520–531.

DOI: 10.1002/ana.21552

Atrial Fibrillation and Stroke Severity: Expanding the Mechanistic Exemplar, Clinical Phenotype, and Goals of Anticoagulant Pharmacotherapy

The relation between left atrial or appendage thrombus and cardioembolic events among individuals with paroxysmal, chronic, and permanent atrial fibrillation is well established and incontrovertible.[1] Thrombus size varies and ranges from a few millimeters to several centimeters.[2] There is a suggestion, based on recent observations, that larger thrombi may bypass the carotid orifice merely as a function of size.[3] In addition, serial transesophageal echocardiograms performed on patients with atrial fibrillation and left atrial thrombus show complete or near-complete resolution after anticoagulant therapy in a majority of instances.[4] These obser-

vations support the importance of preventing thrombus growth as an achievable end point of anticoagulant therapy, permitting intrinsic fibrinolytic systems at the endocardial level to progressively reduce thrombus burden. Thrombus organization, reducing the propensity for embolization, is also an important objective of treatment. Wysokinski and colleagues[5] demonstrated immunohistochemical differences between in situ and embolized thrombi.

In this issue, Ay and colleagues'[6] study validates and extends earlier observations that the level of anticoagulation intensity at presentation correlates with severity of stroke.[7] The investigators found a robust association between international normalized ratio (INR) and acute infarct volume. Their study also raises several fundamental questions. Why should warfarin treatment before a cardioembolic ischemic stroke reduce the volume of injury? Is the protective effect specific to warfarin (vitamin K antagonists) or to the anticoagulant effect itself? If the latter, similar observations would be expected for drugs under development, including factor Xa antagonists such as rivaroxaban and apixiban, and direct thrombin inhibitors such as dabigatran.

Does Vitamin K Antagonist-Based Anticoagulant Therapy Facilitate Thrombus Organization?

Vascular calcification commonly accompanies atherosclerosis involving the coronary arteries, carotid arteries, and aorta, and is caused by deposition of basic calcium phosphate crystals in a highly regulated process.[8] Matrix γ-carboxyglutamic acid (Gla) protein is a vitamin K–dependent protein that inhibits arterial calcification in vivo. Its five Gla residues are formed in a posttranslation carboxylation reaction requiring vitamin K as a cofactor.[9] Vitamin K antagonists increase calcium salt deposition in vascular smooth muscle cell culture by preventing carboxylation and serine phosphorylation of matrix γ-carboxyglutamic acid (Gla) protein.[10] Atrial tissue samples obtained at the time of surgery from patients with atrial fibrillation undergoing a Maze procedure overexpress L-type calcium channel α-1 and sarcoplasmic reticulum Ca^{2+}-ATPase messenger RNA, suggesting that vitamin K antagonism may compound an existing defect in calcium homeostasis in this particular patient population.[11]

The stabilization of intravascular (and endocardial) thrombus is achieved through several highly integrated mechanisms involving coagulation proteases, fibrinolytic proteins, inflammatory cells, and matrix-forming cells, a process collectively referred to as *organization.* A pivotal state in thrombus organization is fibroblast recruitment and their subsequent transformation to collagen-producing myofibroblasts.[12] Because circulating progenitor (stem) cells among patients with atrial fibrillation have a particularly strong propensity to transform into fibroblasts (a response to tissue injury), and because a calcium-rich local environment may facilitate osteogenic cell differentiation, vitamin K antagonists may facilitate thrombus organization. Whether the phenomenon of thrombus stabilization is unique to vitamin K antagonists, by mechanisms summarized previously, or driven predominantly by facilitated organization that could represent a more fundamental and generalizable by-product of limiting thrombus growth can readily be discerned through planned substudies in ongoing, phase III clinical trials.

Does Anticoagulant Therapy Reduce Left Atrial Thrombus Size?

The available information, based on serial echocardiographic studies, suggests strongly that anticoagulant therapy reduces left atrial and appendage thrombus size. Coagulation proteases participate directly in thrombotic, inflammatory, and cellular regulatory processes. Thrombin is a pivotal enzyme in each process; however, factor Xa, alone and as a critical component of the prothrombinase complex (responsible for thrombin generation), is important as well. Indeed, factor Xa provokes tissue factor expression from endothelial cells, damaged endocardial cells, smooth muscle cells, and macrophages.[13]

In addition to attenuating endocardial-based proinflammatory and prothrombotic events, anticoagulant therapy may positively influence blood-borne contributions to thrombus growth and development as well. Tissue factor–containing neutrophils, monocytes, and microparticles that circulate within peripheral blood can be transported to sites of endocardial injury where they contribute directly to thrombus propagation. Systemic anticoagulation, although not directly impacting tissue factor production, may attenuate its activity.[14] It is interesting to note that Wysokinski and colleagues[5] found tissue factor in thrombi to be colocalized with platelets and granulocytes predominantly around the periphery of thrombi and speculated that tissue factor was integral to thrombus propagation.

Does Anticoagulant Therapy Influence Thrombus Architecture and Its Intrinsic Characteristics?

The polymerization of fibrin forms a highly resilient scaffold that imparts unique structural, chemical, and biological diversity. Fibrin ultrastructure, in turn, is influenced greatly by the local conditions in which the conversion from fibrinogen to fibrin, and its subsequent covalent linkage, takes place.[15] Models of fibrin gel formation show that flow shear rate, permeability, and thrombin activity are the key regulators. In vitro clots formed in the presence of thrombin at low concentrations contain thick and widely porous fibrin strands that are highly susceptible to fibrinolysis. The reverse has been observed with thrombin at high con-

centrations, where the developing fibrin strands are thin, tightly packed, and resistant to fibrinolysis.[16]

It is likely that all of these processes contribute to the efficacy of anticoagulant therapy and may help to explain Ay and colleagues'[6] findings. As noted by the authors, the relation between INR and infarct volume was not absolute; for example, several patients with an INR of 2.0 or greater had a proximal pattern indicative of a larger embolus, and several patients with subtherapeutic INR had a distal stroke pattern. Given that thrombus growth or regression is a dynamic process, the duration of warfarin therapy before the embolus would be an important consideration, and these data were not provided. In addition, acute fluctuations in INR are more reflective of changes in factor VII levels (shorter half-life) as opposed to steady-state thrombin levels, and this INR variability would attenuate the relation between INR at presentation and infarct volume. As acknowledged, misclassification of stroke subtype may also have contributed.

From a clinical perspective, Ay and colleagues' study[6] reaffirms the importance of therapeutic anticoagulation (INR of 2.0 or greater) for stroke prevention in atrial fibrillation.[6,17,18] It also highlights the potential tradeoffs inherent to use of lower target INR intensity, that is, range 1.5 to 2.5, as a potential strategy to offset bleeding risk.

Newer anticoagulant drugs with wide therapeutic margins and minimal interference with diet and drugs may supplant the oral vitamin K antagonists. However, the increasing challenge for stroke prevention in atrial fibrillation is the seemingly intractable risk for hemorrhage and ongoing eligibility for anticoagulant drugs, particularly among elderly individuals.[19] As our understanding advances of the complex interplay of basic mechanisms that govern thrombus initiation, propagation, and embolization, hopefully other pivotal targets in these pathways will be identified for interventions devoid of bleeding risk.

Elaine M. Hylek, MD, MPH[1] and
Richard C. Becker, MD[2]

[1]Department of Medicine, Research Unit-Section of General Internal Medicine and Clinical Epidemiology Unit
Boston University School of Medicine, Boston Medical Center
Boston, MA

[2]Divisions of Cardiovascular Medicine and Hematology
Duke University School of Medicine, Duke Cardiovascular Thrombosis Center
Durham, NC

Potential conflict of interest: E.M.H. has received research funding from AstraZeneca and Bristol-Myers Squibb and honoraria from Bayer. R.C.B. has received research funding from AstraZeneca, Bayer, Bristol-Myers Squibb, Johnson & Johnson, Medicines Company, Regado Biosciences, and Schering-Plough Research Institute.

References

1. Parekh A, Jaladi R, Sharma S, et al. Images in cardiovascular medicine. The case of a disappearing left atrial appendage thrombus: direct visualization of left atrial thrombus migration, captured by echocardiography, in a patient with atrial fibrillation, resulting in a stroke. Circulation 2006;114:e513–e514.
2. Thambidorai SK, Murray RD, Parakh K, et al. Utility of transesophageal echocardiography in identification of thrombogenic milieu in patients with atrial fibrillation (an ACUTE ancillary study). Am J Cardiol 2005;96:935–941.
3. McBane RD, Hodge DO, Wysokinski WE. Clinical and echocardiographic measures governing thromboembolism destination in atrial fibrillation. Thromb Haemost 2008;99: 951–955.
4. Tsai LM, Chen JH, Lin LJ, Teng JK. Natural history of left atrial spontaneous echo contrast in nonrheumatic atrial fibrillation. Am J Cardiol 1997;80:897–900.
5. Wysokinski WE, Owen WG, Fass DN, et al. Atrial fibrillation and thrombosis: immunohistochemical differences between in situ and embolized thrombi. J Thromb Haemost 2004;2: 1637–1644.
6. Ay H, Arsava EM, Gungor L, et al. Admission international normalized ratio and acute volume in ischemic stroke. Ann Neurol 2008:64:499–506.
7. 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;349:1019–1026.
8. Demer LL, Tintut Y. Vascular calcification: pathobiology of a multifaceted disease. Circulation 2008;117:2938–2948.
9. Schurgers LJ, Teunissen KJ, Knapen MH, et al. Novel conformation-specific antibodies against matrix γ-carboxyglutamic acid (Gla) protein: undercarboxylated matrix Gla protein as marker for vascular calcification. Arterioscler Thromb Vasc Biol 2005;25:1629–1633.
10. Schurgers LJ, Spronk HM, Skepper JN, et al. Post-translational modifications regulate matrix Gla protein function: importance for inhibition of vascular smooth muscle cell calcification. J Thromb Haemost 2007;5:2503–2511.
11. Adams J, Pepping J. Vitamin K in the treatment and prevention of osteoporosis and arterial calcification. Am J Health Syst Pharm 2005;62:1574–1581.
12. Feigl W, Susani M, Ulrich W, et al. Organisation of experimental thrombosis by blood cells. Evidence of the transformation of mononuclear cells into myofibroblasts and endothelial cells. Virchows Arch A Pathol Anat Histopathol 1985;406: 133–148.
13. Senden NH, Jeunhomme TM, Heemskerk JW, et al. Factor Xa induces cytokine production and expression of adhesion molecules by human umbilical vein endothelial cells. J Immunol 1998;161:4318–4324.
14. Hoffman M, Whinna HC, Monroe DM. Circulating tissue factor accumulates in thrombi, but not in hemostatic plugs. J Thromb Haemost 2006;4:2092–2093.
15. Weisel JW. Structure of fibrin: impact on clot stability. J Thromb Haemost 2007;5(suppl 1):116–124.
16. Wolberg AS. Thrombin generation and fibrin clot structure. Blood Rev 2007;21:131–142.
17. 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: 133–144.

18. Partington SL, Abid S, Teo K, et al. Pre-admission warfarin use in patients with acute ischemic stroke and atrial fibrillation: the appropriate use and barriers to oral anticoagulant therapy. Thromb Res 2007;120:663–669.
19. Hylek EM, Evans-Molina C, Shea C, et al. Major hemorrhage and tolerability of warfarin in the first year of therapy among elderly patients with atrial fibrillation. Circulation 2007;115: 2689–2696.

DOI: 10.1002/ana.21551

Probing Racial and Regional Disparities in Stroke Mortality: Same Problem, Different Solutions?

Beyond its impact as the third leading cause of death in the United States,[1] mortality from stroke uniquely manifests itself in substantive racial and regional disparities.[2,3] For instance, the stroke death rate in African Americans is about twice the rate noted in white counterparts.[2] Regardless of race, people residing in the southeastern 12-state region of the United States known as the Stroke Belt have excess stroke mortality rates compared with the rest of the country.[3] The exact reasons for these disparities remain unclear and somewhat controversial[4]; however, differences in stroke mortality may be explained by a comparatively greater prevalence or severity of vascular risk factors in these disadvantaged groups.[2,4] Probing the nature of these disparities could lead to interventions, thereby ameliorating the overall societal burden of stroke.

Investigating the underlying causes behind these racial and regional stroke mortality disparities would require a fairly large prospective nationwide study of a population adequately representative of these overly burdened groups. The National Institutes of Health–funded study Reasons for Geographic and Racial Differences in Stroke (REGARDS) may be the ideal setting to probe this question because of its large number of participants (almost 30,000), nationwide scope, and oversampling of African Americans, as well as the availability of in-person measures of traditional and novel vascular risk factors.[5] Cross-sectional analyses using validated stroke risk assessment instruments could provide insights into these racial and regional disparities, and could lead to new avenues to bridge these gaps.

In this issue of *Annals of Neurology,* Cushman and colleagues[6] report the 10-year stroke rates predicted by the Framingham Stroke Risk Profile among participants living in the Stroke Buckle (the coastal plain regions of NC, SC, and GA), the Stroke Belt (the remainders of NC, SC, and GA, plus AL, MS, TN, AR, and LA), and the other 40 contiguous states.[6] Overall, the authors found that estimated stroke rates were robustly greater among African Americans compared with their white counterparts, regardless of regional location. Furthermore, among African Americans, these greater predicted stroke rates were noted in the context of significantly greater prevalence of vascular risk factors, including hypertension, increased systolic blood pressure, diabetes, smoking, and left ventricular hypertrophy. On the other hand, and somewhat surprisingly, predicted stroke rates were only marginally greater for residents of the Stroke Belt and Stroke Buckle compared with the rest of the country (10.7 and 10.4 vs 10.1%, respectively), and only two factors, the presence of diabetes and antihypertensive use, were significantly more common in the Stroke Belt compared with the rest of the nation.

The greater rate of predicted stroke and the greater prevalence of vascular risk factors among African Americans compared with whites have also been shown in prior studies.[2] Interventions aimed at modifying vascular risk factors could reduce the occurrence of stroke in African Americans, thereby mitigating the greater stroke mortality rate in this population.[2] Cushman and colleagues'[6] study is also in agreement with earlier data indicating that, despite the impact of greater stroke rates on stroke-related mortality, this difference is probably not large enough to fully account for the substantially greater stroke death rate among African Americans compared with whites. Other factors, such as risk factor severity and socioeconomic status, are likely at play.[2]

The relatively small differences in Framingham Stroke Risk Profile among the geographical regions that Cushman and colleagues[6] studied were unexpected in light of previous reports of a 50% greater stroke mortality rate in the Stroke Belt and Stroke Buckle regions compared with the rest of the nation. If a greater frequency of stroke does contribute substantially to regional stroke mortality differences, what could explain these large differences in regional stroke death rates?

Outcome after an incident stroke can be influenced by a variety of factors, including stroke severity, pathophysiologic mechanism, socioeconomic status, medical comorbidities, and quality of care. The prospective REGARDS study data will provide much needed information about the contributions of stroke severity and subtype, socioeconomic status, and medical comorbidities to these regional stroke mortality disparities. However, about comorbidities, the cross-sectional REGARDS data raise the question whether the presence of established vascular risk factors may contribute to regional

Disparate Stroke Rates on Warfarin Among Contemporaneous Cohorts With Atrial Fibrillation

Potential Insights Into Risk From a Comparative Analysis of SPORTIF III Versus SPORTIF V

Elaine M. Hylek, MD, MPH; Lars Frison, PhD; Lori E. Henault, MPH; Adrienne Cupples, PhD

Background and Purpose—The rate of stroke among warfarin-treated patients in SPORTIF V was approximately half that of patients enrolled in SPORTIF III (1.16%/year versus 2.30%/year). SPORTIF III was an open-label trial comparing ximelagatran with warfarin for stroke prevention in atrial fibrillation. SPORTIF V was a double-blind trial performed in North America. The trial design was otherwise identical. We sought to determine if differences in baseline characteristics, use of potentially risk-modifying medications, or anticoagulation control help to explain the lower risk of stroke among warfarin-treated patients in SPORTIF V.

Methods—Cox regression with stepwise model selection was used to define the covariates independently associated with stroke. Secondary analyses identified covariates with the strongest influence on the study factor (V/III). These covariates were then added to the primary model. Cox regression was used to determine the degree of confounding exerted by these covariates that might help to explain the differences between the trials.

Results—Independent risk factors for stroke on warfarin included prior stroke/transient ischemic attack, coronary artery disease, international normalized ratio, weight, and study. Patients in SPORTIF V were at half the risk as those in SPORTIF III. We found that lower international normalized ratio variability, a higher proportion of prevalent warfarin use, lower systolic blood pressure, high-density lipoprotein, and a greater proportion of statin use among patients in SPORTIF V collectively conferred a lower risk of stroke.

Conclusion—Differences in blood pressure control, international normalized ratio variability, proportion of prevalent warfarin users, statin exposure, and high-density lipoprotein collectively conferred a lower risk of stroke to patients in SPORTIF V. These findings suggest that the different event rates were not due to chance and provide potential insights into stroke risk among warfarin-treated patients with atrial fibrillation. **(*Stroke*. 2008;39:3009-3014.)**

Key Words: atrial fibrillation ■ stroke ■ warfarin

Vitamin K antagonists (eg, warfarin) have been the mainstay of oral anticoagulation for stroke prevention in atrial fibrillation (AF). It is widely recognized that newer antithrombotic agents are needed because of warfarin's narrow therapeutic index and its interaction with food and other drugs.[1] The variability in the dose–response of warfarin mandates frequent monitoring, which also imposes a significant barrier to its use.[2]

The first of these newer agents, an oral direct thrombin inhibitor, ximelagatran, was evaluated in 2 clinical trials enrolling an unprecedented 7329 patients (Stroke Prevention using an ORal Thrombin Inhibitor in atrial Fibrillation [SPORTIF]).[3,4] SPORTIF III was an open-label trial comparing ximelagatran with warfarin conducted in Europe, Asia, Australia, and New Zealand; SPORTIF V was a double-blind trial performed in North America. The trial design was otherwise identical. The primary end point in both trials was all strokes, both ischemic and hemorrhagic, and systemic embolic events. Although potential risk factors were well balanced across treatment arms within each trial, there were significant differences across trials between SPORTIF III and SPORTIF V.[5] The stroke rate among warfarin-treated patients in the North American trial was approximately half that of patients enrolled in SPORTIF III (1.16%/year versus 2.30%/year) despite similar adherence and anticoagulation control.

The objective of our study was to formally assess this difference in event rates among warfarin-treated patients to determine if differences in baseline characteristics, use of potentially risk-modifying concomitant therapies, or variability in control of anticoagulation might help to explain the seemingly lower risk of warfarin-treated patients enrolled in SPORTIF V. Understanding these differences would facili-

Received January 14, 2008; final revision received March 10, 2008; accepted March 31, 2008.

From the Department of Medicine (E.M.H., L.E.H.), Research Unit, Section of General Internal Medicine, Boston University School of Medicine, Boston, Mass; [2]Astra Zeneca R&D (L.F.), Mölndal, Sweden; and the Department of Statistics (A.C.), Boston University, Boston, Mass.

Correspondence Elaine M. Hylek, MD, MPH, Boston University School of Medicine, Research Unit, Section of General Internal Medicine, 801 Massachusetts Avenue, 2nd Floor, Boston, MA 02118. E-mail ehylek@bu.edu

Stroke **is available at http://stroke.ahajournals.org** **DOI: 10.1161/STROKEAHA.108.514935**

tate a more informed interpretation of the 2 trials, provide potential insights into stroke risk among warfarin-treated patients with AF, and highlight important considerations for randomized trials in AF with warfarin as the comparator.

Materials and Methods

The design, patient demographics, and main results of the 2 trials have previously been published.[3,4] The primary end point for both trials included ischemic stroke, hemorrhagic stroke, and systemic embolic events. SPORTIF III enrolled 1703 patients in the warfarin arm. Median follow-up was 17.9 months. In SPORTIF V, 1962 patients were randomized to warfarin and the median follow-up was 20.1 months. In total, 93 primary events (9 primary intracerebral hemorrhages) were confirmed among patients randomized to warfarin in the 2 trials. Outcome assessment was blinded to treatment in both trials.

Variables

Anticoagulation Control (time-in-range, international normalized ratio variability, prevalent warfarin use)

Warfarin dosing and monitoring of the international normalized ratio (INR) were different in the 2 trials. SPORTIF V used a double-blind design with sham INR testing. A standardized point-of-care instrument and uniform thromboplastin reagent were used for the majority of INR measurements. Warfarin dosing was performed by a central laboratory. In the open-label SPORTIF III trial that involved 23 countries, warfarin management was conducted locally by investigators or individual patient physicians. Assays encompassed a variety of methods, reagents, and instruments.

Anticoagulation control within the 2 trials was determined by time-in-range analyses using linear interpolation between INR measurements.[6] Per protocol, INR measurements were mandated at least every 31 days. As a result, gaps in monitoring were rare. Variability in INR was assessed using SD. An additional variable of interest was the proportion of patients taking warfarin at study entry. To the extent that prevalent users represent a warfarin-tolerant population, this survivor bias would be reflected in fewer events overall, both ischemic and hemorrhagic, lower INR variability, and a higher percentage of time in range. The increased INR variability attributable to new warfarin use would heighten the already expected differences due to the use of multiple thromboplastin reagents, instruments, and testing methodologies. Differences in INR variability or time in the therapeutic range may have placed SPORTIF V patients at lower risk for the primary end point.

Potential Risk-Modifying Concomitant Therapy

Differences in exposure to other medications that have the potential to alter stroke risk over time were also evaluated. Ascertainment of concomitant medications in the SPORTIF program was conducted at prespecified points throughout the trial periods. Specific drugs of interest included hydroxymethylglutaryl coenzyme A reductase inhibitors (statins), aspirin, and drugs that inhibit the renin–angiotensin–aldosterone system. Because the majority of patients who were taking these medications at baseline were still taking them at the end of follow-up, we analyzed medication as a binary variable.

Additional Covariates

Other variables of interest that may have contributed to differences in risk included demographic features, known risk factors for stroke (hypertension, heart failure, diabetes, and prior stroke), weight, body mass index, high-density lipoprotein (HDL), low-density lipoprotein, creatinine clearance, and smoking status. Because differences in blood pressure control may have existed between the 2 trials, we also assessed the effect of mean systolic and diastolic blood pressure in addition to analyzing hypertension as a binary variable. Physical examinations that included determination of vital signs were performed at prespecified intervals according to the study protocol: baseline, 1, 3, and 6 months and 6-month intervals throughout the trial period.

Statistical Analyses

The Kolmogorov-Smirnov test was adopted for testing the equality of different empirical distribution functions. Cox regression models with stroke/systemic embolic event as the dependent variable were used to first define the univariate associations of the covariates and outcome among patients taking warfarin in both trials. Stepwise model selection was used for the multivariable Cox regression to define the covariates that were independently associated with stroke and included only those variables from univariate analysis that achieved significance at the 5% level and maintained this level of significance in the presence of other selected covariates in the multivariate model. In the primary univariate analyses, a statistically significant study factor (V/III) was found, ie, patients on warfarin in SPORTIF V were at approximately half the estimated risk for stroke as patients on warfarin in SPORTIF III (hazard ratio=0.541; 95% CI, 0.358 to 0.824). We next sought to define the relationship between covariates and the study factor with bivariate analyses. Using Cox regression, we evaluated whether differences in distributions of risk factors or other variables between the 2 studies could help to explain the difference in event rates. Models with study factor (V/III) plus one additional covariate, one at a time, were used to assess the effect of the covariate on the estimated hazard ratio (0.54) of the study factor. Those factors having the largest impact were then added to the primary multivariable model. Analyses were performed in SAS (version 8.2; SAS Institute, Inc, Cary, NC).

Results

Baseline Characteristics

As previously highlighted in the pooled analysis, there were significant differences in baseline characteristics between SPORTIF V and SPORTIF III.[5] Patients randomized to warfarin in the North American (SPORTIF V) trial were older (mean age, 71.6 years versus 70.1 years) and more had hypertension (Table 1). Fewer patients had a history of stroke (18% versus 24%), and a greater proportion of patients were obese. In SPORTIF V, more patients were taking warfarin at study entry (85% versus 73%), a statin (47% versus 30%), and aspirin (25% versus 17%). Similar proportions of patients were taking angiotensin-converting enzyme inhibitors.

Blood Pressure Control

Although more patients in SPORTIF V had hypertension, blood pressure was controlled to a greater degree in SPORTIF V (Table 2).[7] Mean systolic blood pressure at entry was 132.6 mm Hg versus 139.0 mm Hg among patients in SPORTIF III. Longitudinal assessment of blood pressure revealed that 78% of warfarin-treated patients in SPORTIF V had a mean systolic blood pressure of less than 140 mm Hg compared with 63% of warfarin-treated patients in SPORTIF III ($P<0.001$).

Anticoagulation Control

Despite different methodologies for warfarin dosing and INR measurement, the overall time spent in the therapeutic range, 2.0 to 3.0, was nearly identical in the 2 trials, 68% versus 66% for SPORTIF V versus SPORTIF III and 83% versus 81% for time in the expanded range of 1.8 to 3.2 as has been previously reported.[5] Proportion of time in the INR range <2 was 20% versus 19% and 12% versus 15% for INR >3.

Table 1. Baseline Clinical Characteristics of Warfarin-Treated Patients in SPORTIF III versus SPORTIF V*

	SPORTIF III (N=1703)	SPORTIF V (N=1962)
Variables, % (n)		
Age ≥75 yr	33 (565)	42 (820)†
Female	30 (507)	31 (609)
Prior stroke/TIA	24 (405)	18 (348)†
Hypertension	72 (1230)	81 (1582)†
Heart failure‡	34 (584)	40 (788)†
Diabetes mellitus	22 (377)	25 (483)
Coronary artery disease	40 (675)	48 (944)†
Creatinine clearance <50, mL/min	14 (228)	13 (260)
Vitamin K antagonist at entry	73 (1235)	85 (1661)†
Body weight >100 kg	11 (182)	25 (494)†
Body mass index >30, kg/m^2	31 (520)	40 (790)†
HDL, mmol/L, mean (SD)§	1.27 (0.38)	1.20 (0.39)†
LDL, mmol/L, mean (SD)§	3.16 (0.93)	2.82 (0.86)†
Concomitant medications, % (n)		
Aspirin	17 (290)	25 (482)†
Angiotensin-converting enzyme inhibitor	53 (896)	53 (1041)
Angiotensin-receptor blocker	12 (202)	15 (296)
HMG-CoA reductase inhibitor	30 (504)	47 (928)†
ACEI+HMG-CoA reductase inhibitor	17 (290)	29 (566)†

*Stroke Prevention using an ORal Thrombin Inhibitor in atrial Fibrillation=SPORTIF.

†P<0.001.

‡Left ventricular ejection fraction <40% or symptomatic heart failure.

§Multiply by 39 to convert to mg/dL.

TIA indicates transient ischemic attack; LDL, low-density lipoprotein; HMG-CoA, hydroxymethylglutaryl coenzyme A; ACEI, angiotensin-converting enzyme inhibitor.

However, the proportion of individual patients whose INR was maintained between 2.0 and 3.0 for at least 50% of the time was different in SPORTIF V versus SPORTIF III, 85% versus 79.5% (P<0.001). This finding at least in part reflects the higher proportion of prevalent warfarin users enrolled in the North American trial (Figures 1 and 2). INR variability was lower in SPORTIF V (SD, 0.63 versus 0.72). INR variability was considerably lower among those patients taking warfarin at study entry compared with those newly starting therapy (SD, 0.61 versus 0.85; P<0.001).

Table 2. Distribution of Warfarin-Treated Patients in SPORTIF III and SPORTIF V by Mean Systolic Blood Pressure During the Trial Period (P<0.001)

Mean SBP Category, mm Hg	SPORTIF III (N=1703)	SPORTIF V (N=1962)
84–119.9	214 (12.6%)	410 (20.9%)
120–139.9	865 (50.8%)	1111 (56.6%)
140–159.9	561 (32.9%)	411 (21.0%)
160–179.9	61 (3.6%)	30 (1.5%)
180–191.7	2 (0.1%)	0

SBP indicates systolic blood pressure.

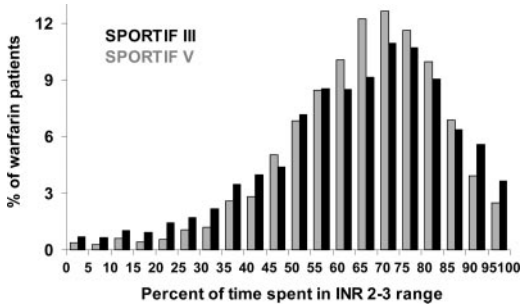


Figure 1. Distribution of patients on warfarin by percentage of time spent in the INR range of 2 to 3. The hypothesis of equal distributions was rejected, P=0.0004, by the Kolmogorov-Smirnov test.

Risk Factors for Stroke and Study Factor Analysis

Among patients taking warfarin, independent risk factors for stroke included prior stroke/transient ischemic attack, coronary artery disease, time in INR range, weight, and study, ie, patients in SPORTIF V were at approximately 40% lower risk than patients in SPORTIF III (Table 3). Covariates identified in bivariate analyses as having the strongest influence on the study factor (V/III) were added to the primary multivariable model. As shown in Table 4, after adjusting for these 5 factors, ie, systolic blood pressure, HDL, statin use, INR variability, and prevalent warfarin use, the effect of the study factor was diminished and no longer statistically significant. This finding suggests that the collective impact (joint confounding) of these covariates helps to explain the disparate stroke event rates experienced within the 2 trials. These factors were sufficiently correlated with the study factor such that after adjusting for them, there was no longer a significant difference between the study groups.

Discussion

Warfarin is highly effective in reducing the risk of ischemic stroke in AF. In the pooled analysis of the first 5 randomized trials, the event rate on placebo was 4.94% compared with 1.91% on oral anticoagulant therapy.[8] In the open-label SPORTIF III trial, the rate of all stroke/systemic embolism was 2.30% on warfarin and 1.16% in SPORTIF V. The difference in event rates on warfarin between the 2 trials has not previously been studied in detail. In this analysis, we found a joint confounding effect of systolic blood pressure,

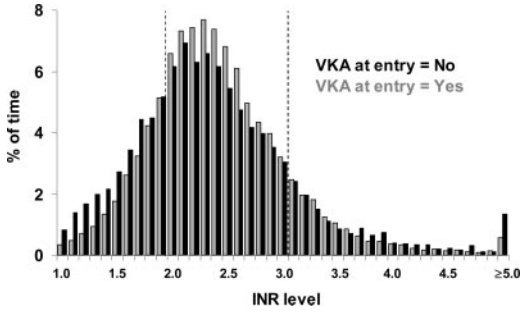


Figure 2. Distributions for the percentage of time by INR level according to warfarin status at entry. The hypothesis of equal distributions was rejected, P<0.0001, by the Kolmogorov-Smirnov test.

Table 3. Independent Predictors for Stroke/Systemic Embolic Event Among Patients Taking Warfarin*

	Univariate		Multivariate	
Variable	HR	(95% CI)	HR	(95% CI)
Study factor V/III	0.54	(0.36–0.82)	0.61	(0.39–0.94)
Prior stroke/TIA	2.56	(1.69–3.89)	2.16	(1.40–3.32)
Coronary artery disease	1.60	(1.06–2.41)	1.75	(1.15–2.65)
Percent time INR >3.0	1.02	(1.00–1.03)	1.02	(1.00–1.03)
Weight	0.20	(0.08–0.50)	0.26	(0.10–0.68)

*Thirty-five covariates were assessed by univariate analysis in addition to the study factor (V/III). Only variables significant at the 5% level, in the presence of other selected variables, were retained in the final model. Variables in italics were significant at the 5% level in the unadjusted analysis: *age ≥75 years*, gender, *previous stroke/TIA*, diabetes, left ventricular dysfunction, *mean SBP at baseline*, *mean SBP during study*, mean DBP at baseline, mean DBP during study, *coronary artery disease*, smoker now, smoker ever, mean heart rate at baseline, mean heart rate during study, *weight*, total cholesterol, HDL, LDL, creatinine clearance, *warfarin at study entry*, *mean warfarin dose*, SD of warfarin dose, body mass index, paroxysmal AF, *aspirin*, class of medication (*beta-blocker*, calcium-channel blocker, angiotensin-converting enzyme inhibitor, angiotensin-receptor blocker, statin), mean INR, *INR SD*, *percent time INR within 2.0 to 3.0*, *percent time INR >3.0*, percent time INR <2.0, *study factor*.

HR indicates hazard ratio; TIA, transient ischemic attack; SBP, systolic blood pressure, DBP, diastolic blood pressure; LDL, low-density lipoprotein.

Table 4. Joint Confounding Effect After Addition of the 5 Covariates Most Strongly Associated With the Study Factor Into the previous Multivariate Analysis*

	Univariate		Multivariate	
Variable	HR	(95% CI)	HR	(95% CI)
Study factor V/III	0.54	(0.36–0.82)	0.77	(0.48–1.23)
Prior stroke/TIA	2.56	(1.69–3.89)	2.32	(1.47–3.64)
CAD	1.60	(1.06–2.41)	1.87	(1.19–2.93)
% time INR >3.0	1.02	(1.00–1.03)	1.01	(1.00–1.03)
Weight	0.20	(0.08–0.50)	0.27	(0.09–0.76)
SBP, mean	1.02	(1.00–1.03)	1.01	(1.00–1.03)
HDL	0.92	(0.56–1.51)	0.78	(0.46–1.35)
Statin therapy	0.73	(0.47–1.13)	0.70	(0.43–1.13)
INR, SD	1.44	(1.09–1.90)	1.30	(0.82–2.06)
Warfarin at entry	0.63	(0.41–0.99)	0.87	(0.52–1.44)

*In a univariate Cox regression, the study factor had HR=0.54 (95% CI, 0.36 to 0.82). In bivariate Cox regression models, the HR (95% CI) for the study factor changed as follows, when each of these 5 top-ranked risk factors/variables was added, one at a time: SBP mean (HR=0.59 [0.38 to 0.89]), HDL (HR=0.57 [0.37 to 0.88]), VKA at entry (HR=0.57 [0.37 to 0.87]), statin use (HR=0.56 [0.37 to 0.86]), INR, SD (HR=0.56 [0.36 to 0.86]). Their individual impact in explaining event rate differences between studies was not sufficient, but their collective impact was.

HR indicates hazard ratio; TIA, transient ischemic attack; CAD, coronary artery disease; SBP, systolic blood pressure; VKA, vitamin K antagonist.

HDL, statin use, INR variability, and prevalent warfarin status that may help to explain the lower event rate among patients taking warfarin in SPORTIF V.

Despite similar aggregate time in range, more individual patients in SPORTIF V maintained the therapeutic range compared with SPORTIF III. SPORTIF V enrolled 85% of patients taking warfarin at study entry compared with 73% in SPORTIF III, and INR variability was less. SPORTIF V used a double-blind design with sham INR testing, a standardized point-of-care instrument, uniform thromboplastin, and a centralized laboratory for dosing of warfarin. These controlled experimental conditions would be difficult to duplicate in real-world care, and in this respect, SPORTIF III more closely reflects clinical practice. The effect of INR intensity on risk of stroke has been well described.[9,10] This was also recently shown in a pooled analysis of warfarin-treated patients in the SPORTIF program.[11] White et al found the rate of stroke/systemic embolism to be 2.10% among patients with <60% time in the therapeutic INR range compared with 1.07% for patients with >75% time in range. In contrast to the early AF trials that established the efficacy of warfarin versus placebo, newer antithrombotic drugs are being compared with warfarin. There is an important survivor bias inherent to prevalent warfarin use that has been underrecognized. New use of warfarin is more closely associated with incident AF, which is a risk-prone period for stroke and hemorrhage.[12–15] Prevalent warfarin use denotes a warfarin-tolerant, lower-risk population.[16] In addition, patients who have been taking warfarin for longer periods of time have a unique training advantage in knowing their own individual triggers for aberrant INR control and have benefited from multiple medical encounters that reinforced healthy behaviors and medication adherence. This frequent interface with medical care may also have helped to optimize other stroke risk factors through improved control of hypertension, diabetes, and heart failure. For all of these reasons, the distinction between warfarin-naive and prevalent warfarin use is important for trials with warfarin as the comparator. Event rates will be influenced by the proportions of these patients enrolled.

Hypertension is a potent risk factor for stroke and its effects persist despite anticoagulant therapy.[9,17] In a pooled analysis of SPORTIF data, Lip et al found the rate of stroke/systemic embolism to increase substantially at mean systolic blood pressures of 140 mm Hg and greater (2.4% versus 1.4% for mean systolic blood pressure <140 mm Hg).[7] In addition to the differences related to warfarin, our data suggest that blood pressure control, statin use, and HDL collectively lowered the risk of patients enrolled in SPORTIF V. This finding is supported by a growing body of evidence affirming the complex endothelial and antithrombotic actions of statin drugs and HDL.[18–26] Evidence also exists for additional benefits of combination therapy and effects of renin–angiotensin–aldosterone inhibitors that are independent of the blood pressure-lowering effects.[27–31] Estimate of risk related to hypertension is obscured by studies that rely on diagnostic codes for hypertension rather than blood pressure measurement, which is essential to discriminate among these different effects.

Distinct from stroke risk factors derived from placebo populations, risk of stroke among individuals taking warfarin provides insight into factors that either influence risk independent of anticoagulation or factors that modulate the efficacy of warfarin. There are comparatively fewer studies of stroke risk among warfarin-treated patients with AF. The prospective study of 364 individuals by Poli and colleagues

was limited by the small number of events (n=21).[17] We found prior stroke to persist as a strong risk factor among warfarin-treated patients as previously reported.[9,17] Coronary artery disease has not consistently been found to be an independent risk factor for stroke in AF after adjusting for other known vascular risk factors.[8,32] However, coronary artery disease is a marker of atherosclerotic burden and an independent predictor of complex aortic plaque, which has been shown to be independently associated with high thromboembolic risk among patients with AF.[33] The protective effect of weight is unexpected. The paradoxical effect of obesity on cardiovascular outcomes has been reported in patients with hypertension and coronary artery disease, patients with heart failure, and in patients treated with early percutaneous coronary intervention for unstable angina/non-ST-segment elevation myocardial infarction.[34–36] Touted mechanisms include detection and treatment bias among obese patients compared with nonobese patients, upregulation of endogenous cannabinoids, and complex effects of inflammatory cytokines and adiponectin.[37–40] Weight may also reflect warfarin dose and warfarin sensitivity as suggested by the pooled analysis that found improved INR control with increasing weight.[11]

Our study has several limitations. The findings were derived from a post hoc analysis of the SPORTIF trials. Residual unmeasured confounding may still exist. We were unable to measure the effect of homocysteine, which has been reported to be a risk factor for stroke among patients with AF taking oral anticoagulants.[17] Given the use of multiple thromboplastins of varying sensitivity, comparison of INR distributions across trials is less than optimal. The most valid comparison would require reassay of plasma at a reference laboratory.

Summary

In this study, we sought to better understand the differences in event rates among warfarin-treated patients with AF enrolled in 2 contemporaneous trials. We found that differences in blood pressure control, INR variability, proportion of prevalent warfarin users, statin exposure, and HDL collectively conferred a lower risk of stroke to patients in SPORTIF V. These findings suggest that the different event rates were not due to chance and provide potential insights into stroke risk among warfarin-treated patients with AF.

Source of Funding

AstraZeneca funded this study and the SPORTIF III and SPORTIF V trials.

Disclosures

E.M.H. has served in an advisory capacity and received research grant support from AstraZeneca and Bristol-Myers Squibb; L.F. is an employee of AstraZeneca.

References

1. Ansell J, Hirsh J, Poller L, Bussey H, Jacobson A, Hylek E. The pharmacology and management of the vitamin K antagonists: the Seventh ACCP Conference on Antithrombotic and Thrombolytic Therapy. *Chest.* 2004;126(suppl):204S–33S.
2. Birman-Deych ERM, Nilasena DS, Gage BF. Use and effectiveness of warfarin in Medicare beneficiaries with atrial fibrillation. *Stroke*. 2006; 37:1070–1074.
3. Executive Steering Committee for the SPORTIF III Investigators. Stroke prevention with the oral direct thrombin inhibitor ximelagatran compared with warfarin in patients with non-valvular atrial fibrillation. *Lancet.* 2003;362:1691–1698.
4. Executive Steering Committee for the SPORTIF V Investigators. Ximelagatran vs warfarin for stroke prevention in patients with nonvalvular atrial fibrillation. *JAMA*. 2005;293:690–698.
5. Diener HC; Executive Steering Committee SPORTIF III and V Investigators. Stroke prevention using the oral direct thrombin inhibitor ximelagatran in patients with non-valvular atrial fibrillation. Pooled analysis from the SPORTIF III and V studies. *Cerebrovasc Dis*. 2006;21: 279–293.
6. Rosendaal FR, Cannegieter SC, van der Meer FJ, Briet E. A method to determine the optimal intensity of oral anticoagulant therapy. *Thrombosis & Haemostasis*. 1993;69:236–239.
7. Lip GY, Frison L, Grind M. Effect of hypertension on anticoagulated patients with atrial fibrillation. *Eur Heart J*. 2007;28:752–759.
8. Atrial Fibrillation Investigators. 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: 1449–1457.
9. 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;335:540–546.
10. SPAF 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;348:633–638.
11. 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 anticoagulation control. *Arch Intern Med*. 2007;167: 239–245.
12. Wang TJ, Massaro JM, Levy D, Vasan RS, Wolf PA, D'Agostino RB, Larson MG, Kannel WB, Benjamin EJ. A risk score for predicting stroke or death in individuals with new-onset atrial fibrillation in the community: the Framingham Heart Study. *JAMA*. 2003;290:1049–1056.
13. Lehto M, Snapinn S, Dickstein K, Swedberg K, Nieminen MS. Prognostic risk of atrial fibrillation in acute myocardial infarction complicated by left ventricular dysfunction: the OPTIMAAL experience. *Eur Heart J*. 2005; 26:350–356.
14. 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;70:725–733.
15. DiMarco JP, Flaker G, Waldo AL, Corley SD, Greene HL, Safford RE, Rosenfeld LE, Mitrani G, Nemeth M; AFFIRM Investigators. Factors affecting bleeding risk during anticoagulant therapy in patients with atrial fibrillation: observations from the AFFIRM Study. *Am Heart J.* 2005; 149:650–656.
16. Hylek EM, Evans-Molina C, Shea C, Henault LE, Regan S. Major hemorrhage and tolerability of warfarin in the first year of therapy among elderly patients with atrial fibrillation. *Circulation*. 2007;115:2689–2696.
17. Poli D, Antonucci E, Cecchi E, Marcucci R, Liotta AA, Cellai AP, Lenti M, Gensini GF, Abbate R, Prisco D. Culprit factors for the failure of well-conducted warfarin therapy to prevent ischemic events in patients with atrial fibrillation: the role of homocysteine. *Stroke*. 2005;36: 2159–2163.
18. Casani L, Sanchez-Gomez S, Vilahur G, Badimon L. Pravastatin reduces thrombogenicity by mechanisms beyond plasma cholesterol lowering. *Thrombosis & Haemostasis*. 2005;94:1035–1041.
19. Fukuta H, Sane DC, Brucks S, Little WC. Statin therapy may be associated with lower mortality in patients with diastolic heart failure: a preliminary report. *Circulation*. 2005;112:357–363.
20. Mineo C, Deguchi H, Griffin JH, Shaul PW. Endothelial and antithrombotic actions of HDL. *Circ Res*. 2006;98:1352–1364.
21. Sanguigni V, Pignatelli P, Lenti L, Ferro D, Bellia A, Carnevale R, Tesauro M, Sorge R, Lauro R, Violi F. Short-term treatment with atorvastatin reduces platelet CD40 ligand and thrombin generation in hypercholesterolemic patients. *Circulation*. 2005;111:412–419.
22. Sola S, Mir MQS, Lerakis S, Tandon N, Khan BV. Atorvastatin improves left ventricular systolic function and serum markers of inflammation in nonischemic heart failure. *J Am Coll Cardiol*. 2006;47:332–337.

23. Wadham C, Albanese N, Roberts J, Wang L, Bagley CJ, Gamble JR, Rye KA, Barter PJ, Vadas MA, Xia P. High-density lipoproteins neutralize C-reactive protein proinflammatory activity. *Circulation*. 2004;109: 2116–2122.
24. Haendeler JH, Zeiher AM, Dimmeler S. Antioxidant effects of statins via S-nitrosylation and activation of thioredoxin in endothelial cells. *Circulation*. 2004;110:856–861.
25. Lee T-S, Chang C-C, Zhu Y, Shyy J. Simvastatin induces heme oxygenase-1. *Circulation*. 2004;110:1296–1302.
26. Eichinger S, Pecheniuk NM, Hron G, Deguchi H, Schemper M, Kyrle PA, Griffin JH. High density lipoprotein and the risk of recurrent venous thromboembolism. *Circulation.* 2007;115:1609–1614.
27. Koh KK, Quon MJ, Han SH, Chung WJ, Ahn JY, Seo YH, Kang MH, Ahn TH, Choi IS, Shin EK. Additive beneficial effects of losartan combined with simvastatin in the treatment of hypercholesterolemic, hypertensive patients. *Circulation*. 2004;110:3687–3692.
28. Arima H, Hart RG, Colman S, Chalmers J, Anderson C, Rodgers A, Woodward M, MacMahon S, Neal B. Perindopril-based blood pressure-lowering reduces major vascular events in patients with atrial fibrillation and prior stroke or transient ischemic attack. *Stroke*. 2005;36:2164–2169.
29. Olsson LG, Swedberg K, Ducharme A, Granger CB, Michelson EL, McMurray JJ, Puu M, Yusuf S, Pfeffer MA. Atrial fibrillation and risk of clinical events in chronic heart failure with and without left ventricular systolic dysfunction: results From the Candesartan in Heart failure-Assessment of Reduction in Mortality and morbidity (CHARM) Program. *J Am Coll Cardiol*. 2006;47:1997–2004.
30. Wachtell K, Hornestam B, Lehto M, Slotwiner DJ, Gerdts E, Olsen MH, Aurup P, Dahlof B, Ibsen H, Julius S, Kjeldsen SE, Lindholm LH, Nieminen MS, Rokkedal J, Devereux RB. Cardiovascular morbidity and mortality in hypertensive patients with a history of atrial fibrillation: the Losartan Intervention for End Point Reduction in Hypertension (LIFE) Study. *J Am Coll Cardiol*. 2005;45:705–711.
31. Schmieder RE, Hilgers KF, Schlaich MP, Schmidt BM. Renin–angiotensin system and cardiovascular risk. *Lancet*. 2007;369:1208–1219.
32. Watson T, Shanstila E, Lip GYH. Management of atrial fibrillation: an overview of the NICE guidance on AF management. *Br J Cardiol*. 2007;14:23–28.
33. Zabalgoitia M, Halperin JL, Pearce LA, Blackshear JL, Asinger RW, Hart RG. Transesophageal echocardiographic correlates of clinical risk of thromboembolism in nonvalvular atrial fibrillation. Stroke Prevention in Atrial Fibrillation III Investigators. *J Am Coll Cardiol*. 1998;31: 1622–1626.
34. Uretsky S, Messerli FH, Bangalore S, Champion A, Cooper-DeHoff RM, Zhou Q, Pepine CJ. Obesity paradox in patients with hypertension and coronary artery disease. *Am J Med*. 2007;120:863–870.
35. Horwich TB, Fonarow GC, Hamilton MA, MacLellan WR, Woo MA, Tillisch JH. The relationship between obesity and mortality in patients with heart failure. *J Am Coll Cardiol*. 2001;38:789–795.
36. Buettner HJ, Mueller C, Gick M, Ferenc M, Allgeier J, Comberg T, Werner KD, Schindler C, Neumann F. The impact of obesity on mortality in UA/non-ST-segment elevation myocardial infarction. *Eur Heart J*. 2007;28:1694–1701.
37. Wannamethee SG, Whincup PH, Lennon L, Sattar N. Circulating adiponectin levels and mortality in elderly men with and without cardiovascular disease and heart failure. *Arch Intern Med*. 2007;167:1510–1517.
38. Laughlin GA, Barrett-Connor E, May S, Langenberg C. Association of adiponectin with coronary heart disease and mortality: the Rancho Bernardo Study. *Am J Epidemiol*. 2007;165:164–174.
39. Feldman AM, Combes A, Wagner D, Kadakomi T, Kubota T, Li YY, McTiernan C. The role of tumor necrosis factor in the pathophysiology of heart failure. *J Am Coll Cardiol*. 2000;35:537–544.
40. Steffens S, Veillard NR, Arnaud C, Pelli G, Burger F, Staub C, Karsak M, Zimmer A, Frossard JL, Mach F. Low dose oral cannabinoid therapy reduces progression of atherosclerosis in mice. *Nature*. 2005;434: 782–786.

Europace (2008) **10**, 3–5
doi:10.1093/europace/eum258

EDITORIAL

Methodological considerations for interpretation of rates of major haemorrhage in studies of anticoagulant therapy for atrial fibrillation

Elaine M. Hylek[1]* and Jared W. Magnani[2]

[1]*Clinical Epidemiology Unit and Section of General Internal Medicine-Research Unit, Boston University School of Medicine, Boston, MA, USA; and* [2]*Cardiovascular Medicine Division, Boston University School of Medicine, Boston, MA, USA*

Online publish-ahead-of-print 26 November 2007

This editorial refers to 'Anticoagulation for elderly patients with atrial fibrillation: not to be neglected' by J. S. Taggar and G. Y. H. Lip, on page 1

Major haemorrhage is a significant concern for elderly patients with atrial fibrillation (AF) taking anticoagulant therapy. Two recently published studies have demonstrated conflicting results with regard to the incidence of major haemorrhage on warfarin in this age group. The Birmingham Atrial Fibrillation Treatment of the Aged Study[1] (BAFTA) randomized 973 patients with AF aged 75 years or more to warfarin or aspirin. During the 2.7 year mean follow-up, the incidence of major haemorrhage was nearly identical in the two study arms (1.9% per year for warfarin vs. 2.2% for those taking aspirin). In contrast, a single centre, inception cohort study identified a significantly greater incidence of major haemorrhage in elderly patients initiating warfarin for stroke prevention in AF:[2] 7.2% per year in a population with mean age 77 years, increased to 13.1% per year among those aged 80 and older.

How can such disparate rates be reconciled? The answers are found in the key elements of study design which guide the interpretation of event rates. These studies contrasting results demonstrate the importance of assessing study design, patient population (with particular regard to selection bias), length of observation, definition of major haemorrhage, specific anticoagulant, quality of anticoagulation control, and use of other risk-modifying therapy.

Study design

Randomized controlled trials are the most valid assessment of an intervention. Although often criticized for the highly selected nature of the study population, a trial's first priority is to ensure internal validity to facilitate an objective assessment of efficacy. Anticipated drop-out rates should be low and medication adherence rates should be high. Translating trial results into real-world practice can be challenging because the patient populations are often considerably different (external validity). Observational studies provide important insight into clinical practice, but these designs are also subject to bias. A source of bias that affects both randomized trials and observational studies is enrolment of prevalent users of warfarin opposed to those newly starting therapy. Bleeding events are more likely to occur early in the treatment course, and new users are often sicker at baseline and therefore more likely to stop therapy.[3] Patients who remain on warfarin (prevalent users) are the 'survivors' and are a lower risk group compared with those individuals who 'dropped out'. Trials and non-inception cohorts that predominantly enrol prevalent warfarin users will have lower event rates.

In BAFTA, 40% of subjects were taking warfarin at study entry; 100% of patients enrolled in the inception cohort study were newly starting warfarin. The latter study found the risk of haemorrhage was significantly increased during the first 90 days of therapy (IRR 3.31, 95% CI 1.51–7.25). In addition, bleeding and early warfarin termination occurred more often among patients with a $CHADS_2$ score of ≥ 3. This study demonstrated that higher risk patients have a disproportionate burden of adverse events, and that these events occur early following the initiation of warfarin. Thus, BAFTA may have been vulnerable to enrolling a population of patients who were not 'warfarin naïve'. This higher proportion of warfarin prevalent subjects likely influenced the different major haemorrhage incident rates in these populations and emphasized the need for new-user designs in randomized controlled trials of antithrombotic therapy. Overall there were too few events in BAFTA to meaningfully assess this difference.

The opinions expressed in this article are not necessarily those of the Editors of *Europace*, the European Heart Rhythm Association or the European Society of Cardiology.

*Corresponding author. Tel: +1 617 414 3743; fax: +1 617 414 4676.
E-mail address: ehylek@bu.edu

Patient population

As is true of any potentially hazardous medication, the initial selection bias related to 'candidacy' for the drug is very difficult to overcome. This decision is often subjective and frequently based on relative rather than absolute contraindications to anticoagulant therapy. Trials often further restrict enrolment of patients perceived to be at the lowest risk for haemorrhage. The setting in which patients are identified also has critical implications: the period immediately post-hospital discharge is more risk-laden, for instance, than recruitment from routine care in an ambulatory setting. Patients identified from a hospitalization will likely also have a greater degree of comorbidities.

The patient populations enrolled in the two studies were distinctly different. Inclusion in BAFTA was restricted to patients for whom there was clinical uncertainty as to whether aspirin or warfarin should be used. This criterion therefore favoured enrolment of patients at lower risk of stroke and lower risk of haemorrhage. It is challenging to assert equipoise in aspirin vs. warfarin for patients with heart failure or high $CHADS_2$ scores. Given the 50% likelihood of being randomized to warfarin, it also seems unlikely that these patients were deemed to be at high risk of haemorrhage. Exclusion criteria in BAFTA included: major haemorrhage in previous 5 years, intracranial haemorrhage, endoscopically proven peptic ulcer disease in the previous year, oesophageal varices, or surgery within the past 3 months. In contrast, by principle of study design, the inception cohort study had none of these exclusion criteria. Patients were deemed warfarin candidates independent of the study and placed on this anticoagulation regimen, regardless if they had any of the exclusion criteria articulated by the BAFTA investigators. In point of fact, 10% of patients had a history of gastrointestinal or other bleed. BAFTA enrolled 100% of its participants from the ambulatory setting, while the inception cohort study identified a third of its patients at hospital discharge. Twenty-eight percent of patients enrolled in BAFTA had a $CHADS_2$ score of ≥ 3; this is in comparison with the 44% of patients in the cohort study aged 80 years or greater.

Length of observation

Because adverse events are more likely to occur early among new users of anticoagulant therapy, longer-term follow-up disproportionately reflects the experience of the warfarin 'survivors'. The number of hemorrhagic events decreases over time while the person-year denominator continues to be enriched by prevalent users of warfarin.

BAFTA's mean follow up was 2.7 years. However, the inception cohort was designed purposefully to truncate observation at 1 year of warfarin therapy. Had the exposure period been extended, the 'surviving' patients would have continued to contribute to the person-years in the denominator resulting in a lower reported rate of major haemorrhage.

Definition of major haemorrhage

Broad, less restrictive definitions of haemorrhage serve to inflate event rates. Reporting of aggregate bleeding rates (major plus minor) tends to deter use of warfarin in clinical practice. This is particularly problematic in AF given the increased disability and mortality related to stroke in AF. The inception cohort study used the definition of major haemorrhage promulgated by the International Society of Haemostasis and Thrombosis: fatal bleeding or symptomatic bleeding in a critical area or organ, such as intracranial, intraspinal, intraocular, retroperitoneal, intraarticular or pericardial, or intramuscular with compartment syndrome, and/or bleeding causing a fall in haemoglobin level of 20 g/L (1.24 mmol/L) or more, or leading to transfusion of two units of packed red blood cells.[4]

Therefore, the reported rate of 7.2% for major haemorrhage is not due to inclusion of minor events.

Characteristics of the anticoagulant drug

Both studies exclusively used warfarin and experienced nearly identical time above the therapeutic range (BAFTA 14% vs. 13%). Thus, the difference in bleeding rates was not attributable to different drug pharmacokinetics or suboptimal anticoagulation control. Adherence to warfarin in BAFTA was stated to be 67% which may have biased towards lower rates of bleeding in the intention to treat analysis. The investigators acknowledged this possibility, but noted no difference in the on-treatment analyses.

Use of other risk-modifying therapy

It is unclear in BAFTA if patients randomized to warfarin had access to over-the-counter aspirin. The hazards of dual therapy are now being more fully realized as this combination is being prescribed with increasing frequency to older patients.[5,6] In the uncontrolled, unselected observational cohort study, 40% of patients were also taking aspirin for either primary or secondary prevention of cardiovascular disease. This higher than expected prevalence of dual therapy has also been found in other patient populations.[7] Twelve of the 26 major bleeding events in the cohort study occurred on combination therapy.

Summary comments

These two studies are in marked contrast to one another in terms of study design, enrolment, patient characteristics and outcomes. BAFTA enrolled patients with a choice of anticoagulant therapies and then randomized them to warfarin or aspirin. In contrast, the inception cohort study identified a sicker population with more extensive comorbidities who had been placed on warfarin and then followed for the first year. If anything, the BAFTA population reflects the outcome of a healthier population. It would be interesting to know the rate of haemorrhage among the excluded patients in the BAFTA trial whose physicians deemed warfarin to be indicated. These patients are likely more comparable with those patients enrolled in the inception cohort study.

The higher rates of major haemorrhage in the inception cohort study were driven by the unselected nature of the study population, focus on the initiation period, higher proportion of patients identified at hospital discharge, restriction to 1-year follow-up, and high prevalence of concomitant use of aspirin. The cohort study highlights the clinical complexity of patients with AF encountered in routine practice and the need for aggressive strategies to optimize

the benefit of anticoagulant therapy for stroke prevention in this vulnerable patient population.

Conflict of interest: none declared.

References

1. Mant J, Hobbs FD, Fletcher K, Roalfe A, Fitzmaurice D, Lip GY *et al*. BAFTA investigators; Midland Research Practices Network. 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**:460–1.
2. Hylek EM, Evans-Molina C, Shea C, Henault LE, Regan S. Major hemorrhage and tolerability of warfarin in the first year of therapy among elderly patients with atrial fibrillation. *Circulation* 2007;**115**:2689–96.
3. Ray WA. Evaluating medication effects outside of clinical trials: new-user designs. *Am J Epidemiol* 2003;**158**:915–20.
4. Schulman S, Kearon C. On behalf of the Subcommittee on Control of Anticoagulation of the Scientific Standardization Committee of the International Society on Thrombosis and Haemostasis. Definition of major bleeding in clinical investigations of antihemostatic medicinal products in non-surgical patients. *J Thromb Haemost* 2005;**3**:692–4.
5. Eikelboom JW, Hirsh J. Combined antiplatelet and anticoagulant therapy: clinical benefits and risks. *J Thromb Haemost* 2007;**5**(Suppl. 1): 255–63.
6. 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**:117–24.
7. Johnson SG, Witt DM, Eddy TR, Delate T. Warfarin and antiplatelet combination use among commercially insured patients enrolled in an anticoagulation management service. *Chest* 2007;**131**:1500–7.

JACC: CARDIOVASCULAR INTERVENTIONS
VOL. 1, NO. 1, 2008

ISSN 1936-8798/08/$34.00
PUBLISHED BY ELSEVIER
DOI: 10.1016/j.jcin.2007.11.005

EDITORIAL COMMENT

Dual Antiplatelet and Oral Anticoagulant Therapy

Increasing Use and Precautions for a Hazardous Combination*

Elaine M. Hylek, MD, MPH,† David E. Solarz, MD‡

Boston, Massachusetts

The prevalence of atrial fibrillation (AF) in the U.S. is projected to reach nearly 8 million individuals by 2020 (1). Stroke prevention in AF is the most common indication for warfarin. Approximately one-third of these individuals will also have coronary artery disease. Overall, the number of warfarin prescriptions has increased from 21 million in 1998 to nearly 31 million in 2004 (2). This surge in warfarin use has paralleled an increase in aspirin use for primary and secondary prevention of cardiovascular disease. The increased incidence of major hemorrhage, particularly among the older patient population, is at least in part attributable to the increased use of combination antithrombotic therapy (3). This risk was also recently highlighted in the Warfarin Antiplatelet Vascular Evaluation trial that randomized patients with peripheral arterial disease to antiplatelet or combination therapy (4).

See page 56

In this issue of *JACC: Cardiovascular Interventions*, Rogacka et al. (5) report their findings from 127 consecutive patients discharged on aspirin, a thienopyridine (clopidogrel or ticlopidine), and warfarin after coronary stent implantation. Atrial fibrillation was the indication for warfarin in 59% of patients. The mean exposure to combination therapy was 5.6 ± 4.6 months. Of the 127 patients, 6 experienced a major hemorrhage, 4 of which were intracranial hemorrhages (ICH), and 3 were fatal. Most events occurred within the first month.

This study importantly adds to the growing body of evidence documenting the hemorrhagic risk of combined dual antiplatelet and warfarin therapy. The hazards of triple therapy after percutaneous coronary intervention were first highlighted in a retrospective analysis of 65 patients discharged after coronary stenting on aspirin, clopidogrel, and warfarin. Six patients (9.2%; 95% confidence interval [CI] 3.5 to 19.0) experienced a bleeding complication; 2 met criteria for major hemorrhage (6). Khurram et al. (7) subsequently published a retrospective study of 107 patients. Major bleeding occurred in 7 patients (6.6%); the hazard ratio of triple therapy was 5.4 compared with dual antiplatelet therapy alone. Similar findings were reported from Finland. Among 185 patients treated for a mean of 4 months, 18 (8.2%) sustained a major hemorrhage, including 3 ICH (2 related to trauma) (8).

Limitations of these studies include small sample size, retrospective design, lack of international normalized ratio (INR) data, and small number of events that prohibits meaningful assessment of risk. In the study by Rogacka et al. (5), the distribution of hemorrhage raises additional concern. Four of the 6 major bleeds were ICH. Information on blood pressure was not provided. The disproportionate number of ICH suggests incomplete ascertainment of major extracranial hemorrhage and a subsequent under-estimate of the aggregate bleeding rate. It is also important to note that although the authors attempt to assess differences in hemorrhage with different types of stents, in the absence of randomized data, this is problematic because of confounding by indication (i.e., patients at higher risk of hemorrhage might have preferentially received bare-metal stents to minimize exposure to triple therapy).

The risk of upper gastrointestinal bleeding or perforated peptic ulcer with aspirin use increases exponentially with age (9). Aspirin in combination with clopidogrel increases this risk more than 3-fold (adjusted rate ratio [RR] 3.9, 95% CI 2.8 to 5.5), and aspirin plus warfarin increases this risk more than 6-fold (adjusted RR 6.5, 95% CI 4.3 to 9.9) (10). The number needed to harm with antiplatelet therapy is 33 among high-risk female patients with a history of complicated ulcer who are ≥80 years of age and 17 for the comparable risk stratum in men. Nonsteroidal anti-inflammatory drugs (NSAIDs) in combination with aspirin greatly magnify this risk (9).

Potential Strategies to Minimize Hemorrhagic Complications

Consistent risk factors for major hemorrhage on warfarin include older age, anticoagulation intensity, early course of therapy, prior bleed, and concomitant antiplatelet therapy (11,12). Additional risk factors for ICH include prior stroke and hypertension. The clinical challenge is how best to navigate optimal prevention of thrombosis while minimizing serious bleeding consequences. Combination antiplatelet therapy is less effective than warfarin for stroke preven-

*Editorials published in *JACC: Cardiovascular Interventions* reflect the views of the authors and do not necessarily represent the views of *JACC: Cardiovascular Interventions* or the American College of Cardiology.

From the †Clinical Epidemiology Unit and Research Unit–Section of General Internal Medicine, and the ‡Department of Medicine, Boston University School of Medicine, Boston, Massachusetts. Dr. Hylek has received research funding from AstraZeneca and Bristol-Myers Squibb and has served in an advisory capacity to Bristol-Myers Squibb.

Table 1. Potential Strategies to Minimize Major Hemorrhage Among Patients on Triple Therapy After Coronary Stent Implantation

1. Vigilant INR monitoring in the first 4 weeks, especially among patients newly starting warfarin or antiplatelet therapy
2. Judicious use of "bridging therapy" with heparin (e.g., highest-risk mechanical prosthetic heart valve, venous thromboembolism within 3 months)
3. Improved risk stratification for warfarin use in patients with AF (CHADS2 score ≥2)
4. Increased awareness of the most potent risk factors for erratic INR control: decompensated heart failure, enteral feeding, erratic dietary vitamin K intake, amiodarone therapy, chemotherapy, protracted new use of high-dose acetaminophen
5. Attention to blood pressure control with goal <130/80 mm Hg (15)
6. Prophylactic proton-pump inhibition for patients with peptic ulcer disease
7. Eradication of *H. pylori* in patients with peptic ulcer disease and uninvestigated dyspepsia (17)
8. Explicit warnings regarding use of over-the-counter NSAIDs and aspirin-containing compounds
9. Physical therapy/safety evaluation before discharge to minimize fall risk
10. There is insufficient evidence to support lower INR target intensities; patients and their caregivers need to be cognizant of the trade-offs inherent to this strategy

AF = atrial fibrillation; CHADS2 = Congestive heart failure, Hypertension, Age ≥75 years, Diabetes mellitus, prior Stroke/transient ischemic attack; INR = international normalized ratio; NSAIDs = nonsteroidal anti-inflammatory drugs.

tion in AF (13). Improved risk stratification in AF is a logical first step. Current AF guidelines recommend warfarin for CHADS2 scores of ≥2 (Congestive heart failure = 1 point; Hypertension = 1 point; Age ≥75 years = 1 point; Diabetes mellitus = 1 point; prior Stroke/transient ischemic attack = 2 points) and aspirin or warfarin for a CHADS2 score of 1 point (14). This point is emphasized by the current study in which 2 of the patients with ICH had CHADS2 scores = 0 and 1.

It is unclear to what degree hemorrhage on triple therapy is attributable to suboptimal control of warfarin. Proven strategies to minimize anticoagulant-related bleeding should be aggressively implemented, and interventions to ameliorate the hazards of antiplatelet therapy should be instituted (Table 1) (15). Prophylactic treatment with the proton-pump inhibitor (PPI) lansoprazole coupled with eradication of *H. pylori* infection was shown to significantly reduce the risk of recurrence of ulcer complications associated with aspirin (16,17). Patients were treated with 100 mg aspirin and lansoprazole 30 mg daily after confirmed *H. pylori* eradication. The risk of recurrence was 1.6% in the PPI group (95% CI 0% to 9%) and 14.8% in the placebo group (95% CI 7% to 26%); adjusted hazard ratio 9.6 (95% CI 1.2 to 76.1). The number needed to treat was only 7.6 to prevent 1 bleeding complication.

Future Directions

It is possible that newer anticoagulant drugs with shorter half-lives and a wider therapeutic window might be safer to use in combination with antiplatelet therapy. The balance between antithrombotic potency and safety will demand improved risk stratification particularly among individuals ≥75 years of age as highlighted by the recent trial that compared clopidogrel with prasugrel for the prevention of in-stent thrombosis (18). The mean age of trial participants was 61 years, and 13% were ≥75 years of age. Insights into the mechanisms of enhanced gastrointestinal susceptibility to antithrombotic therapy particularly in the elderly patient population warrant further study (19). Further data, preferentially from randomized trials, are needed to better assess the effects of staggered therapy, single versus dual antiplatelet therapy and warfarin, and optimal duration of antiplatelet therapy among different risk subgroups (20).

Reprint requests and correspondence: Dr. Elaine M. Hylek, Boston University School of Medicine, Clinical Epidemiology Unit and Research Unit–Section of General Internal Medicine, 91 East Concord Street, Suite 200, Boston, Massachusetts 02118. E-mail: ehylek@bu.edu.

REFERENCES

1. Miyasaka Y, Barnes ME, Gersh BJ, et al. Secular trends in incidence of atrial fibrillation in Olmstead County, Minnesota, 1980–2000, and implications on the projections for future prevalence. Circulation 2006;114:119–25.
2. Wysowski DK, Nourjah P, Swartz L. Bleeding complications with warfarin use. Arch Intern Med 2007;167:1414–9.
3. Hylek EM, Evans-Molina C, Shea C, Henault LE, Regan S. Major hemorrhage and tolerability of warfarin in the first year of therapy among elderly patients with atrial fibrillation. Circulation 2007;115:2689–96.
4. The Warfarin Antiplatelet Vascular Evaluation Trial Investigators. Oral anticoagulant and antiplatelet therapy and peripheral arterial disease. N Engl J Med 2007;357:217–27.
5. Rogacka R, Chieffo A, Michev I, et al. Dual antiplatelet therapy after percutaneous coronary intervention with stent implantation in patients taking chronic oral anticoagulation. J Am Coll Cardiol Intv 2008;1:56–61.
6. Orford JL, Fasseas P, Melby S, et al. Safety and efficacy of aspirin, clopidogrel, and warfarin after coronary stent placement in patients with an indication for anticoagulation. Am Heart J 2004;147:463–7.
7. Khurram Z, Chou E, Minutello R, et al. Combination therapy with aspirin, clopidogrel and warfarin following coronary stenting is associated with a significant risk of bleeding. J Invasive Cardiol 2006;18:162–4.
8. Karjalainen PP, Porela P, Ylitalo A, et al. Safety and efficacy of combined antiplatelet-warfarin therapy after coronary stenting. Eur Heart J 2007;28:726–32.
9. Hernández-Díaz S, García Rodríguez LA. Cardioprotective aspirin users and their excess risk of upper gastrointestinal complications. BMC Med 2006;4:22.
10. Delaney JA, Opatrny L, Brophy JM, Suissa S. Drug-drug interactions between antithrombotic medications and the risk of gastrointestinal bleeding. CMAJ 2007;177:347–51.
11. Ansell J, Hirsh J, Poller L, Bussey H, Jacobson A, Hylek EM. The pharmacology and management of the vitamin K antagonists: the Seventh ACCP Conference on Antithrombotic and Thrombolytic Therapy. Chest 2004;126:204S–33S.
12. Levine MN, Raskob G, Beyth RJ, Kearon C, Schulman S. Hemorrhagic complications of anticoagulant treatment. Chest 2004;126:287S–310S.
13. The ACTIVE Writing Group on Behalf 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). Lancet 2006;367:1903–12.

JACC: CARDIOVASCULAR INTERVENTIONS, VOL. 1, NO. 1, 2008
FEBRUARY 2008:62–64

14. Fuster V, Ryden LE, Cannom DS, et al. ACC/AHA/ESC 2006 guidelines for the management of patients with atrial fibrillation—executive summary: a report of the American College of Cardiology/American Heart Association Task Force and the European Society of Cardiology Committee on Practice Guidelines. J Am Coll Cardiol 2006;48:854–906.
15. Hart RG, Tonarelli SB, Pearce LA. Avoiding central nervous system bleeding during antithrombotic therapy: Recent data and ideas. Stroke 2005;36:1588–93.
16. Lai KC, Lam SK, Chu KM, et al. Lansoprazole for the prevention of recurrences of ulcer complications from long-term low-dose aspirin use. N Engl J Med 2002;346:2033–8.
17. Chey WD, Wong BC, for the Practice Parameters Committee of the American College of Gastroenterology. American College of Gastroenterology guideline on the management of Helicobacter pylori infection. Am J Gastroenterol 2007;102:1808–25.
18. Wiviott SD, Braunwald E, McCabe CH, et al., the TRITON–TIMI 38 Investigators. Prasugrel versus clopidogrel in patients with acute coronary syndromes. N Engl J Med 2007;357:2001–15.
19. Lanas A, Garcia-Rodriquez LA, Arroyo MT, et al. Effect of antisecretory drugs and nitrates on the risk of ulcer bleeding associated with nonsteroidal anti-inflammatory drugs, antiplatelet agents, and anticoagulants. Am J Gastroenterol 2007;102:507–15.
20. Nguyen MC, Lim YL, Walton A, et al., for the GRACE Investigators. Combining warfarin and antiplatelet therapy after coronary stenting in the Global Registry of Acute Coronary Events: is it safe and effective to use just one antiplatelet agent? Eur Heart J 2007;28:1717–22.

RESEARCH LETTER

The Representation of Women on the Editorial Boards of Major Medical Journals: A 35-Year Perspective

In their recent landmark report on gender bias in science and engineering, the National Academy of Sciences and Institute of Medicine called for "reasonable representation of women on editorial boards and in other significant leadership positions."[1(pS7)] Indeed, membership on the editorial board of a major medical journal is a highly visible, prestigious appointment and affords one the opportunity to have a substantial impact on the nature of the published scholarly discourse of academic medicine. Despite the increased entry of women into the medical profession over the past several decades, with women constituting half of the current medical school class,[2] some have suggested that women may remain substantially underrepresented in senior leadership positions, including editorial positions at biomedical journals.[3]

For editorial comment see page 449

This study examines the gender distribution of editors-in-chief and members of the editorial boards of 16 major medical journals over time to determine the extent to which women have achieved these influential positions within the academic medical community.

Methods. We focused on 16 prominent biomedical journals that publish original research, selected after consideration of published journal impact factors,[4,5] citation half-life, and comments solicited from medical school faculty members regarding the long-term prestige and importance of the various journals in their fields. We included 5 major, English-language, general medical journals published in the United States, United Kingdom, and Canada—*The Journal of the American Medical Association* (*JAMA*), *The New England Journal of Medicine* (*NEJM*), *The Lancet* (*Lancet*), *British Medical Journal* (*BMJ*), and *Canadian Medical Association Journal* (*CMAJ*)—that publish clinically oriented research of interest to practitioners of all medical specialties. We also selected 6 journals specific to 4 major clinical specialty fields: internal medicine (*Annals of Internal Medicine* [*Ann Intern Med*] and *Archives of Internal Medicine* [*Arch Intern Med*]), surgery (*Annals of Surgery* [*Ann Surg*]), obstetrics and gynecology (*Obstetrics and Gynecology* [*Obstet Gynecol*]), and pediatrics (*The Journal of Pediatrics* [*J Pediatr*] and *Pediatrics*). In addition, we examined 5 journals that report research in the experimental biomedical sciences: *Cell*, *Nature Medicine* (*Nature Med*), *Science*, *The Journal of Experimental Medicine* (*J Exp Med*), and *The Journal of Clinical Investigation* (*J Clin Invest*).

The lists of editorial board members and names of editors-in-chief published in the first issues of the years 1970, 1975, 1980, 1985, 1990, 1995, 2000, and 2005 were examined for each journal to determine number and gender. Gender was determined by initial inspection of first name; for cases in which gender was not certain, attempts were made to discern gender through Internet searches using the Google (Google Inc, Mountain View, California) search engine or by communication with the journal's editorial offices.

Other editors (such as senior editors, deputy editors, and assistant editors) and editorial staff members (such as editorial assistants, copy editors, and statistical consultants) were not included in the analyses that are presented here. Gender information was collected for the senior editorial staff of each journal in each of the years studied, and these data are available to interested readers on request to the corresponding author. Given the heterogeneity in titles, qualifications, and duties for editorial staff positions across the different journals, it was not possible to summarize these data with an acceptable degree of clarity or consistency in this report.

For certain journals, the group of biomedically trained professionals appearing to serve the role of editorial board was not called "editorial board" but rather held a different title (eg, International Advisory Board and Advisory Editors). In these cases, the specific title of the individuals analyzed is listed in **Table 1**. Individuals who served as editors-in-chief or editorial board members in multiple years were counted in each year in which their names were listed.

The tabulated data were stored in an Excel database (Microsoft Corporation, Redmond, Washington) and analyzed using SAS version 9.1 statistical software (SAS Institute Inc, Cary, North Carolina) to determine the gender distribution of members of the editorial board for each journal. Percentages were calculated from only those names with gender determined; names for which gender could not be identified were excluded from all analyses. Because the size of the editorial boards varied considerably from journal to journal, an analysis in which the percentage of women was averaged from each journal (to weight equally each journal) was also performed. Reported *P* values pertain to the significance of trends over time in these data.

Results. Gender was determined for 3218 of the 3237 names (99.4%) of editorial board members listed in the selected journals in the years studied. Gender was determined for 115 of 118 editors-in-chief (97.5%) listed.

Table 1 presents the detailed gender distribution of editorial board members and editors-in-chief. Overall, 371 (11.5%) of the editorial board members were female. This figure rose from 1.4% in 1970 to 16.0% in 2005. Significant trends of increased female representation were evident for a number of the analyzed journals that had editorial boards over the 35-year period (Table 1). However,

Table 1. Gender Distribution of Editors and Editorial Board Members of 16 Prominent Biomedical Journals

Journal	1970	1975	1980	1985	1990	1995	2000	2005	*P* Value
JAMA									
Total editorial board members, No.	None listed	27	24	25	24	25	26	25	
Female, No. (%)[a]	None listed	1 (4)	0	5 (20)	3 (13)	4 (17)	3 (12)	4 (16)	.09
Gender of editor-in-chief	M	M	M	M	M	M	F	F	
NEJM									
Total editorial board members, No.	15	15	15	15	24	26	24	19	
Female, No. (%)[a]	1 (7)	0	0	0	1 (4)	3 (12)	3 (13)	2 (11)	.07
Gender of editor-in-chief	M	M	M	M	M	M	F	M	
Lancet									
Total "international advisory board" members, No.	None listed	None listed	None listed	None listed	None listed	25	24	24	
Female, No. (%)[a]	None listed	None listed	None listed	None listed	None listed	2 (9)	3 (13)	10 (42)	.005
Gender of editor-in-chief	M	None listed	None listed	None listed	M	Unknown	M	M	
BMJ[b]									
Total editorial board members, No.	None listed	None listed	None listed	None listed	None listed	12	63	None listed	
Female, No. (%)[a]	None listed	None listed	None listed	None listed	None listed	3 (27)	19 (31)	None listed	.82
Gender of editor-in-chief	M	M	None listed	None listed	M	M	M	None listed	
CMAJ									
Total editorial board members, No.	None listed	None listed	None listed	None listed	None listed	None listed	22	18	
Female, No. (%)[a]	None listed	None listed	None listed	None listed	None listed	None listed	7 (33)	3 (18)	.27
Gender of editor-in-chief	M	M	Unknown	M	M	M	M	M	
Ann Intern Med									
Total editorial board members, No.	18	18	24	18	18	18	18	16	
Female, No. (%)[a]	0	1 (6)	2 (8)	2 (11)	2 (11)	5 (28)	7 (39)	5 (31)	<.001
Gender of editor-in-chief	M	M	M	M	M	M	M	M	
Arch Intern Med									
Total editorial board members, No.	10	12	15	23	15	12	8	14	
Female, No. (%)[a]	0	0	0	0	0	2 (17)	2 (25)	6 (43)	<.001
Gender of editor-in-chief	M	M	M	M	M	M	M	M	
J Pediatr									
Total editorial board members, No.	14	15	31	25	22	24	23	19	
Female, No. (%)[a]	0	1 (7)	4 (13)	4 (16)	6 (27)	6 (25)	3 (13)	3 (16)	.12
Gender of editor-in-chief	M	M	M	M	M	M	M	M	
Pediatrics									
Total editorial board members, No.	28	28	26	22	25	27	27	23	
Female, No. (%)[a]	1 (4)	1 (4)	4 (15)	4 (18)	3 (12)	6 (22)	6 (22)	5 (22)	.007
Gender of editor-in-chief	M	M	M	M	M	M	M	M	

(*continued*)

women remained in the minority on all editorial boards studied, and markedly so for some. Fewer than 15% of the editorial board members of the *NEJM* and *Cell* in each year studied, and fewer than 10% of the members of the editorial board of *Ann Surg* in each year were women.

Table 2 presents the mean percentage of women among editorial board members, with each journal weighted equally. Overall, this percentage rose from 1% in 1970 to 21% in 2005. Table 2 also details the mean percentage of women by journal type. In 2005, the mean percentage of women was 22% in the general medical journals, 25% in the clinical specialty medical journals, and 15% in the biomedical science journals. These differences were not statistically significant ($P=.89$ for the comparison of the percentage female in the general medical journals vs the basic science journals and $P=.27$ for the percentage female in the specialty medical vs basic science journals).

In the year 2000, all 5 general medical journals listed editorial boards; these were compared by the journal's national origin. The representation of women on the editorial boards of the 3 general medical journals published in the United Kingdom and Canada differed significantly from the proportion of women on the editorial boards of the 2 general medical journals published in the United States in that year. Of the 50 editorial board members for the 2 American general journals (*JAMA* and *NEJM*), 6 (12.0%) were female, and 29 (27.1%) of the 107 editorial board members of the Canadian (*CMAJ*) and British (*The Lancet* and *BMJ*) journals were female. This difference was statistically significant ($P=.04$).

Overall, 8 (7.0%) of the editors-in-chief listed were female (7 individuals, with 1 individual serving during 2 of the years studied). The journals ever having female editors-in-chief in the studied years were *JAMA*, *NEJM*, *Cell*, *J Exp Med*, and *Nature Med*. Journals having female editors-in-chief were not significantly more likely to have female editorial board members during those years than journals with male editors-in-chief ($P=.45$).

Comment. This study demonstrates a substantial increase in women's representation on the editorial boards of a number of major medical journals over the past 3 decades and the appointment of women to the position of

Table 1. Gender Distribution of Editors and Editorial Board Members of 16 Prominent Biomedical Journals (cont)

Journal	1970	1975	1980	1985	1990	1995	2000	2005	*P* Value
Ann Surg									
Total "editorial and advisory board" members, No.	26	21	27	25	24	37	46	58	.08
Female, No. (%)[a]	0	0	1 (4)	1 (4)	1 (4)	1 (3)	3 (7)	4 (7)	
Gender of editor-in-chief	M	M	M	M	M	M	M	M	
Obstet Gynecol									
Total editorial board members, No.	8	8	9	12	13	13	15	16	.03
Female, No. (%)[a]	0	2 (25)	0	1 (8)	2 (15)	3 (23)	4 (27)	5 (31)	
Gender of editor-in-chief	M	M	M	M	M	M	M	M	
Cell									
Total "associate editors," No.	Journal not started	25	28	46	61	61	69	64	.17
Female, No. (%)[a]	Journal not started	2 (8)	0	3 (7)	3 (5)	5 (8)	5 (7)	7 (11)	
Gender of editor-in-chief	Journal not started	M	M	M	M	M	F	F	
Science									
Total editorial board members, No.[c]	14	14	9	14	None listed	19	80	117	.20
Female, No. (%)[a]	0	2 (14)	2 (22)	2 (14)	None listed	3 (16)	11 (14)	22 (19)	
Gender of editor-in-chief	M	M	M	M	M	M	M	M	
J Clin Invest[d]									
Total editorial committee members and consulting editors, No.	None listed	30	35	34	31	140	180	217	.10
Female, No. (%)[a]	None listed	2 (7)	0	1 (3)	0	22 (16)	11 (6)	22 (10)	
Gender of editor-in-chief	None listed	Unknown	M	M	M	M	M	M	
J Exp Med									
Total "advisory editors," No.	16	17	29	38	40	98	99	101	<.001
Female, No. (%)[a]	0	0	2 (7)	2 (5)	2 (5)	12 (12)	14 (14)	18 (18)	
Gender of editor-in-chief[e]	M	M	M	M	M	M	M	F	
Nature and *Nature Med*[f]									
Gender of editor-in-chief	None listed	M	None listed	M	M	F	F	M	

Abbreviations: *Ann Intern Med, Annals of Internal Medicine; Ann Surg, Annals of Surgery; Arch Intern Med, Archives of Internal Medicine; BMJ, British Medical Journal; CMAJ, Canadian Medical Association Journal; J Clin Invest, The Journal of Clinical Investigation; J Exp Med, The Journal of Experimental Medicine; J Pediatr, The Journal of Pediatrics; NEJM, The New England Journal of Medicine; Nature Med, Nature Medicine; Obstet Gynecol, Obstetrics and Gynecology.*

[a] Number is reported as number female/total number with gender determined.

[b] In 1995, *BMJ* listed a group of "editorial advisors" only, and this is the group analyzed in the Table. In 2000, it listed an "editorial board" (33 members, of whom 10 were female) and "editorial advisors" (30 individuals, of whom 9 were female and 1 unknown); these groups were combined for this.

[c] In 2000 and 2005, *Science* did not list an editorial board, but it did list a "board of reviewing editors." Therefore, for 1970 to 1995, we considered the "editorial board" lists, and in 2000 and 2005, the list of the "board of reviewing editors."

[d] From 1975 to 1990, *J Clin Invest* listed an "editorial committee." In 1995, it listed both an "editorial committee" of 36 individuals, of whom 12 were female, as well as a list of 104 "consulting editors," of whom 10 were female and 3 were unknown. In 2000 and 2005, it listed only "consulting editors." In these analyses, in 1975-1990, the editorial committee was analyzed; in 1995, both groups were combined; and in 2000 to 2005, consulting editors were analyzed.

[e] *J Exp Med* did not list a specific individual as editor-in-chief in the years 1970, 1975, 1980, 1985, 1990, or 1995. The journal did, however, list a small group of "editors" (4 in 1970, 5 in 1975, 4 in 1980, 3 in 1985, 5 in 1990, and 6 in 1995). These individuals were all male; therefore, the gender of the editor-in-chief has been reported and analyzed as "male" for each of these years. In 2000 and 2005, an "executive editor" was listed separately, and this individual was analyzed as the editor-in-chief.

[f] The editors-in-chief of *Nature* are considered here, until the time that *Nature Med* was spun off (1995, 2000, and 2005). *Nature* does not employ an editorial board of any sort (written communication, Juan Carlos Lopez, Editor-in-Chief, *Nature Med*, June 26, 2006).

Table 2. Mean Percentage of Female Editorial Board Members by Journal Type and Year

	All		Category					
			Biomedical Science		General Medical		Clinical Specialty	
Year	No. of Journals	Mean % Female	No. of Journals	Mean % Female	No. of Journals	Mean % Female	No. of Journals	Mean % Female
1970	9	1	2	0	1	7	6	1
1975	12	6	4	7	2	2	6	7
1980	12	6	4	7	2	0	6	7
1985	12	9	4	7	2	10	6	10
1990	11	9	3	3	2	8	6	12
1995	14	17	4	13	4	16	6	20
2000	15	18	4	10	5	20	6	22
2005	14	21	4	15	4	22	6	25

editor-in-chief of several prominent journals. Nevertheless, the majority of editors-in-chief of the journals studied herein continue to be men, and women's representation on the editorial boards of certain journals remains low.

Previous studies have also explored the gender distribution of journal editors in years past.[6,7] Wilkes and Kravitz[8] conducted a survey study a decade ago, finding that 96% of 221 responding editors-in-chief of US and Canadian medical journals were men. Furthermore, they found that 96% of respondents reported having "a great deal of control over" original scientific articles, commentaries, and editorials, leading them to speculate that "homogeneity and concentration of power may have contributed to certain unintended biases in publication."[8(p447-448)] However, a detailed study of the gender of editors, reviewers, and corresponding authors of manuscripts submitted to *JAMA* in 1991 revealed no apparent effect of gender differences on the final outcome of the review process or acceptance for publication.[9]

In any case, it is clear that the positions of editor-in-chief and editorial board member are prestigious appointments that afford the opportunity to guide the underlying direction of a journal, with the potential for consequent impact on the nature of the research emphasized in the medical academic community. These individuals make critical decisions and policies that govern the dissemination of scientific information, such as rules protecting the public from information that has not yet been reviewed by the academic community, policies regarding conflicts of interest on the part of researchers, and policies regarding advertisements.[8] Therefore, it is important to consider the evolving composition of these appointments and the underlying forces shaping that distribution over time.

Our results show that in 2005, 16% of the editorial board members of the journals studied were women. The mean percentage of women for that year (with each journal weighted equally) was 21%. The greatest challenge in interpreting these findings lies in determining what the level of women's participation ought to have been. If the proportion of medical students who are women has only recently reached parity, one cannot expect the senior ranks (from which editorial board members are drawn) to approach an equal gender distribution until substantial time has passed. This has been dubbed the "pipeline phenomenon." Unfortunately, several well-designed studies have suggested that the low representation of women observed in senior positions in academic medicine is due to more than a simple pipeline effect alone.[10] For example, a detailed cohort study by Nonnemaker[11] has suggested that women are not achieving senior status as quickly as would be expected. Tesch and colleagues[12] have also found that promotions in rank appear to be achieved more slowly by women faculty members. A number of explanations for these findings have been proposed,[13-23] and future research should seek to understand more fully the barriers faced by women in their ascent to the upper echelons of academic medicine.

Also worthy of further exploration is the possibility that women's participation may actually improve the ability of senior professional bodies such as editorial boards to serve society. As one editor-in-chief noted in his discussion of the creation of an editorial board for his journal, "The main purpose for developing a board is to help us to prepare the journal for the next century. The members will ensure a steady flow of new ideas from a wide range of specialties, countries, and disciplines."[24(p52)] Just as diversity in specialty, country, and discipline may be important to the ability of an editorial board to serve its mission, so may be gender diversity. If women have different perspectives and different life experiences than their male colleagues, new ideas and approaches might emerge from their participation in these groups, and the critical decisions they make might be improved. Moreover, women's participation in these visible, influential editorial positions may encourage the continued, vital participation of women in medicine more generally by providing visible role models for younger women. We believe these are important subjects for further investigation.

The primary limitation of this study lies in the small numbers that compose the denominator for many of the individual journal-years. Therefore, it is important to consider these data primarily in the aggregate, so that small, chance fluctuations (such as the hypothetical stepping down of several female editors just prior to a year considered in the study) would have only minimal effects on the percentage of women observed. We also wish to emphasize that the comparison between US and British/Canadian journals reported in the "Results" section was restricted to the one year in which the 5 general medical journals all reported editorial boards, and therefore we view this as an exploratory rather than definitive analysis, worthy of further investigation if possible in future years.

As Kuhn[25(p210)] notes, "Scientific knowledge, like language, is intrinsically the common property of a group or else nothing at all. To understand it we shall need to know the special characteristics of the groups that create and use it." While this study documents heartening trends of increased female representation on the editorial boards of a number of the prominent journals that shape our scientific knowledge, it also suggests a need for a greater understanding of barriers that may have impeded (and may continue to impede) even greater participation by women in these critical leadership positions.

Reshma Jagsi, MD, DPhil
Nancy J. Tarbell, MD
Lori E. Henault, MPH
Yuchiao Chang, PhD
Elaine M. Hylek, MD, MPH

Correspondence: Dr Jagsi, Department of Radiation Oncology, University of Michigan Hospital, 1500 E Medical Center Dr, Room B2C490, Ann Arbor, MI 48109-5010 (rjagsi@med.umich.edu).

Author Contributions: Drs Jagsi and Hylek had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. *Study concept and design:* Jagsi and Hylek. *Acquisition of data:* Jagsi, Tarbell, Henault, and Hylek. *Analysis and interpretation of data:* Jagsi, Tarbell, Chang, and Hylek. *Drafting of the manuscript:* Jagsi and Hylek. *Critical revision of the manuscript for important intellectual content:* Jagsi, Tarbell, Henault, Chang, and Hylek. *Statisti-*

cal analysis: Henault and Chang. *Obtained funding:* Jagsi, Tarbell, and Hylek. *Administrative, technical, and material support:* Tarbell, Henault, and Hylek. *Study supervision:* Tarbell and Hylek.

Financial Disclosure: None reported.

Funding/Support: This work was supported by a grant from the American Medical Association Women Physicians' Congress Joan F. Giambalvo Memorial Scholarship Fund.

Role of the Sponsor: The funding source played no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; or preparation, review, or approval of the manuscript.

Additional Contributions: Rebecca Starr and Brenda Vega contributed in data collection.

1. National Academy of Sciences, National Academy of Engineering, and Institute of Medicine, Committee on Maximizing the Potential of Women in Academic Science and Engineering. *Beyond Bias and Barriers: Fulfilling the Potential of Women in Academic Science and Engineering*. Washington, DC: National Academies Press; 2006.
2. American Association of Medical Colleges. *Women in US Academic Medicine: Statistics and Medical School Benchmarking 2004-2005*. Washington, DC: American Association of Medical Colleges; 2005:Table 1.
3. Hatfield C, Ostbye T, Sori C. Sex of editor in medical journals [letter]. *Lancet*. 1995;345(8950):662.
4. Thomson Institute for Scientific Information. *Journal Citation Reports*. Philadelphia, PA: ISI Press; 2002.
5. Garfield E. The history and meaning of the journal impact factor. *JAMA*. 2006; 295(1):90-93.
6. Dickersin K, Fredman L, Flegal K, Scott JD, Crawley B. Is there a sex bias in choosing editors? epidemiology journals as an example. *JAMA*. 1998;280 (3):260-264.
7. Misakian AL, Williard R, Bero L. Gender representation in the editorial and peer review process. Paper presented at: Third International Congress on Peer Review in Biomedical Publication; September 20, 1997; Prague, Czechoslovakia.
8. Wilkes MS, Kravitz RL. Policies, practices, and attitudes of North American medical journal editors. *J Gen Intern Med*. 1995;10(8):443-450.
9. Gilbert JR, Williams ES, Lundberg GD. Is there gender bias in *JAMA*'s peer review process? *JAMA*. 1994;272(2):139-142.
10. Jagsi R, Guancial E, Cooper-Worobey C, et al. The gender gap in academic medical publication: a 35-year perspective. *N Engl J Med*. 2006;355(3): 281-287.
11. Nonnemaker L. Women physicians in academic medicine: new insights from cohort studies. *N Engl J Med*. 2000;342(6):399-405.
12. Tesch BJ, Wood HM, Helwig AL, Nattinger AB. Promotion of women physicians in academic medicine: glass ceiling or sticky floor? *JAMA*. 1995; 273(13):1022-1025.
13. Yedidia MJ, Bickel J. Why aren't there more women leaders in academic medicine? the views of clinical department chairs. *Acad Med*. 2001;76(5):453-465.
14. Levey BA, Gentile NO, Jolly HP, Beaty HN, Levey GS. Comparing research activities of women and men faculty in departments of internal medicine. *Acad Med*. 1990;65(2):102-106.
15. Committee on Women Faculty. A study on the status of women faculty in science at MIT. 1999 http://web.mit.edu/fnl/women/women.html. Accessed December 18, 2007.
16. Kaplan SH, Sullivan LM, Dukes KA, Phillips CF, Kelch RP, Schaller JG. Sex differences in academic advancement—results of a national study of pediatricians. *N Engl J Med*. 1996;335(17):1282-1289.
17. Carr P, Friedman RH, Moskowitz MA, Kazis LE, Weed HG. Research, academic rank, and compensation of women and men faculty in academic general internal medicine. *J Gen Intern Med*. 1992;7(4):418-423.
18. Buckley LM, Sanders K, Shih M, Hampton CL. Attitudes of clinical faculty about career progress, career success and recognition, and commitment to academic medicine. *Arch Intern Med*. 2000;160(17):2625-2629.
19. Okin SM. *Justice, Gender, and the Family*. New York, NY: Basic Books; 1989.
20. Carr PL, Ash AS, Friedman RH, et al. Relation of family responsibilities and gender to the productivity and career satisfaction of medical faculty. *Ann Intern Med*. 1998;129(7):532-538.
21. Grosz B, Dulac C, Dymecki S, et al. Report from the Harvard Task Force on Women in Science and Engineering, May 2005. http://www.news.harvard .edu/gazette/daily/2005/05/wise.pdf. Accessed September 23, 2005.
22. Barres BA. Does gender matter? *Nature*. 2006;442(7099):133-136.
23. Hamel MB, Ingelfinger JR, Phimister E, Solomon C. Women in academic medicine—progress and challenges. *N Engl J Med*. 2006;355(3):310-311.
24. Smith R. BMJ creates an editorial board. *BMJ*. 1997;314(7073):52-58.
25. Kuhn TS. *The Structure of Scientific Revolutions*. Chicago, IL: University of Chicago Press; 1970.

COMMENTS AND OPINIONS

Low-Dose Rosiglitazone in Patients With Type 2 Diabetes Mellitus Requiring Insulin Therapy

We would like to raise the following issues regarding the interesting article by Hollander et al.[1] The high dropout rates (176 of the 630 randomized subjects) reported in the study may have reduced power and comparability and thus limit the validity of the results.[2] Furthermore, including sample size estimation in the manuscript would facilitate assessment of the impact of follow-up losses on the study results.

The investigators must describe cardiovascular adverse events in detail, in view of recent reports showing that rosiglitazone is associated with higher risk of myocardial infarction.[3] In addition, confounder analysis should be performed to evaluate if the increase in cardiovascular events reported in the study could be attributed to the high prevalence of risk factors in the subjects as reported by the authors.

Although diet and exercise are significant confounders of diabetes control,[4,5] these variables were neither measured nor their effects on outcomes analyzed. Insulin dose was left to discretion of the investigators, which could confound diabetes control and further limit validity.

Cesar A. Lopez, MD
Balavenkatesh Kanna, MD, MPH

Correspondence: Dr Lopez, Department of Internal Medicine, Lincoln Medical and Health Center, 234 E 149th St, Bronx, NY 10451 (doctorcesarlopez@yahoo.com).

1. Hollander P, Yu D, Chou HS. Low-dose rosiglitazone in patients with insulin-requiring type 2 diabetes. *Arch Intern Med*. 2007;167(12):1284-1290.
2. Wittes J. Sample size calculations for randomized controlled trials. *Epidemiol Rev*. 2002;24(1):39-53.
3. Nissen SE, Wolski K. Effect of rosiglitazone on the risk of myocardial infarction and death from cardiovascular causes. *N Engl J Med*. 2007;356(24):2457-2471.
4. Sigal RJ, Kenny GP, Wasserman DH, Castaneda-Sceppa C, White RD. Physical activity/exercise and type 2 diabetes: a consensus statement from the American Diabetes Association. *Diabetes Care*. 2006;29(6):1433-1438.
5. Bantle JP, Wylie-Rosett J, Albright AL. Nutrition recommendations and interventions for diabetes—2006: a position statement of the American Diabetes Association. *Diabetes Care*. 2006;29(9):2140-2157.

Neither Freedom nor Autonomy Without Beneficence

Varma and Wendler[1] recently focused on the medical ethical challenge represented by the treatment for people lacking advanced directives or designated surrogates. They argue that it is fundamental to give them the same level of respect afforded those with surrogates. The clear identification of instruments to help physicians make treatment decisions consistent with the patient's preferences is mandatory. They propose a "population-based treatment indicator," a computer-based tool that should be able to unravel a patient's choice

J Pediatr Adolesc Gynecol (2008) 21:195–200

Original Studies

Pre-Pregnancy Body Mass Index among Pregnant Adolescents: Gestational Weight Gain and Long-Term Post Partum Weight Retention

Natalie Pierre Joseph, MD, MPH[1], Kodjo Bossou Hunkali, MD, MPH[2], Bridget Wilson, BS, Elisha Morgan, BS[1], Meagan Cross[1], and Karen M. Freund, MD, MPH[3]

[1]Department of Pediatrics, Boston University School of Medicine; [2]Boston University School of Public Health; [3]Women's Health Unit, Evans Department of Medicine and Women's Health Interdisciplinary Research Center, Boston University School of Medicine, Boston, Massachusetts, USA

Abstract. *Study Objective:* To determine the relationship between adolescents' pre-pregnancy body mass index (BMI), and gestational weight gain and postpartum weight retention.

Design: We review the medical records of adolescents participating in a prospective cohort study on comprehensive health care and parenting education to determine pre-pregnancy BMI, gestational weight gain, and postpartum weight retention at one year.

Setting: Urban academic hospital clinic.

Participants: 102 pregnant adolescents aged 15–21 years.

Main Outcomes: Gestational weight gain and weight retention at one year postpartum.

Results and Conclusions: Fifty-two (51%) adolescent women had normal pre-pregnancy BMI according to the Institute of Medicine classification. Adolescent women with normal (36.5%) and high pre-pregnancy BMI (66.5%) were more likely than women with low pre-pregnancy BMI (26.5%) to exceed recommended gestational weight gain. Adolescent women who exceeded recommended weight gain retained significantly more weight at 1 year postpartum than women with weight gain within or below the recommendation. In a linear regression model that controlled for age, smoking, pregnancy complication, and post partum contraceptive use, pre-pregnancy BMI and gestational weight gain were the strongest predictors of postpartum weight retention at 1 year.

A normal to high pre-pregnancy BMI and excessive gestational weight gain are important predictors of postpartum weight retention in adolescents. These two predictors must be monitored systematically with the aim of preventing postpartum obesity and its associated diseases among this population.

Address correspondence to: Natalie Pierre Joseph, MD, MPH, Boston Medical Center, 850 Harrison Ave, ACC-5, Adolescent Center, Boston MA, 02118; E-mail: Natalie.pierrejoseph@bmc.org

Key Words. Pregnancy in adolescents—Body Mass Index—Gestational weight gain—Postpartum weight retention—Obesity

Introduction

Four in 10 adolescent girls get pregnant at least once before the age of 20 in the United States, leading to nearly 900,000 teenage pregnancies each year.[1–3] Among adolescents, there is some evidence those with excessive maternal weight gain has a greater tendency to retain excess weight postpartum,[4] placing them at risk of becoming obese. There is a higher rate of obesity among adolescents from lower income groups as opposed to higher income groups.[5,6] The identification of groups of people at risk of becoming obese is one of the Healthy People 2010 objectives for prevention and research. However, little is known about weight retention among adolescent women following delivery and its relationship with weight gain during pregnancy.

The purpose of this study was to examine the correlates of excessive maternal weight gain among adolescent mothers using a clinic based sample of pregnant adolescents from an urban, academic teaching hospital. We examine the association between adolescent women's pre-pregnancy and postpartum weights 12 months following delivery with attention to how much they gained during pregnancy. It also examines the implications of gaining weight according to the latest Institute of Medicine recommendations for women's postpartum weight. Recognizing that larger weight gains may be associated with

1083-3188/08/$22.00
doi:10.1016/j.jpag.2007.08.006

subsequent obesity, the Institute of Medicine has identified this as an area where further research is needed.[7]

Materials and Methods

This study is part of a larger prospective cohort study of the effectiveness of comprehensive health care and parenting education to pregnant and parenting adolescents on adolescent parenting, the development of self-sufficiency and maternal and child health. The current study addresses pre-pregnancy body mass index (BMI), gestational weight gain and postpartum weight retention among adolescents.

Subjects

The study subjects were parenting adolescents within the larger prospective cohort study. Subjects were recruited at their child's primary care visit at the Teen and Tot clinic between January 2002 to January 2005 as part of a demonstration care project. Eligible subjects were 21 years or less at enrollment, with a singleton infant at delivery or whose child was 18 months or younger at recruitment. The subjects lived in neighborhoods adjacent to the Boston Medical Center hospitals. The Institutional Review Boards of Boston Medical Center approved the study.

Methods and Procedure

The medical records were reviewed quarterly from before pregnancy until the adolescent mother was three years postpartum. For the purpose of this study, records from the electronic medical record of the adolescent mother and child were reviewed for information up until one year postpartum. The following were reviewed from the medical record: the pre-pregnancy weight (the last weight before a positive pregnancy test was confirmed); pre-pregnancy BMI (calculated as weight in kilograms divided by the square of height in meters); weight gain during pregnancy was calculated from pre-pregnancy weight and last pregnancy weight recorded before delivery either at the last prenatal visit or at delivery. Postpartum weight retention was calculated as the difference between the gestational weight gain(weight at delivery or 2−6 weeks postpartum − prepregancy weight) and the postpartum weight loss (weight at delivery/or 2−6 weeks postpartum − weight at 1 year postpartum) Other data abstracted from the medical record included: age at menarche, gravidity, parity, smoking history during pregnancy, and contraception use at 6 months postpartum. Young women's actual weight gain during pregnancy was classified according to the Institute Of Medicine's guidelines/ reclassification (Table 1).

Table 1. Institute of Medicine Recommendations for Weight Gain in Pregnancy

Initial Body Mass Index	IOM Recommended Gestational Weight Gain (Pound)
<19 (low)	28−40
19.8−26.0 (normal)	25−35
26.1−29 (high)	15−25
>29.0 (obese)	At least 15*

*The obese pregnant women are to gain at least 15 pounds like the women with high BMI. but with no upper limit of weight gain set for obese women.

Statistical Evaluation

Data were analyzed by testing bivariate relationships among study measures. Main outcome variables were: gestational weight gain and weight retention at one year postpartum. To evaluate for independent associations of pregnancy weight gain in teen mothers on these two continuous dependent variables, linear regression analyses were conducted to control for potential confounders. The adjusted potential confounding variables used in the regression to predict gestational weight gain are: age of mother at delivery, pre-pregnancy BMI, smoking during pregnancy, age at menarche, contraception use at six months and pregnancy complications The model to predict postpartum weight retention in teen mothers used a linear regression model with all the above confounding variables, and in addition, included pregnancy weight gain. A *P*-value less than 0.05 was considered statistically significant.

We utilized the Institute of Medicine's 1990 weight gain guidelines for pregnant women, the first guidelines to adjust for pre-pregnancy weight and height. The BMI has been found to be a better indicator of nutritional status than weight alone.[7] Based on these guidelines, ''underweight'' women (BMI < 19.8) should be advised to gain 28−40 pounds during pregnancy; ''overweight'' women (BMI = 26.1 to 29.0) should be advised to gain 15−25 pounds; and ''normal'' weight women (BMI = 19.8 to 26.0) 25 to 35 pounds. Obese women (BMI > 29.0)should be advised to gain at least 15 pounds with no upper limit stated.

Results

One hundred twelve adolescent women were eligible for this analysis. Ten were excluded because of missing information on weight retention at 1 year postpartum, leaving a final sample size was 102. Descriptive maternal characteristics of the sample are shown in Table 2.

At enrollment, the mean age of adolescent mothers in the study was 15 ± 1.5 years. Ninety-eight percent

Table 2. Description of Sample (N = 102)

Maternal Characteristic	N (%)
Age	
Mean age ± SD years	15 (14-20) ± 1.5
<15	19 (19%)
15-17	44 (43%)
17-20	39 (38%)
Gravida	
1	72 (71%)
2+	30 (29%)
Para	
0	100 (98%)
1	2 (2%)
Gestation (wks)	
<37	13 (13%)
≥37	89 (87%)
Maternal race or ethnicity	
African American	88 (84%)
Hispanic	13 (12%)
White	1 (1%)
Other	3 (3%)
Contraceptive use postpartum (at ≤ 6 month)	
Yes	80 (79%)
No	21 (21%)

were primiparous. Eighty-four percent of the sample was African American, 12% were Latina, 1% was non-Hispanic white, and 3% were of other ethnicities. Seventy-five percent of the sample described their ''culture of origin'' as African American, 3% Puerto Rican, 3% Haitian, 1% Dominican, 10% described it as ''other.''

The majority of the adolescents were between ages 15−17 years (43%), followed by the young adult age group of 17−20 years (38%). Seventy-one percent of these young women were pregnant for the first time, 98% were giving birth for the first time, and of those who delivered, 87% (89) of the adolescent mothers delivered at full term. Fifty-one percent (52) of adolescent women had a normal pre-pregnancy BMI. Thirty percent met the Institute of Medicine (IOM) weight gain recommendation, and 36% exceeded the IOM recommendation (Table 3).

Bivariate analyses showing the relationship of pre-pregnancy BMI to IOM weight gain recommendations are displayed in Table 3. Young women with normal (37%) and high pre-pregnancy BMI (67%) were most likely to have exceeded IOM recommended weight gain , followed by young with normal BMI (37%), women with low BMIs (27%) and obese young women (25%). Young women whose gestational weight gain was above the IOM recommendation retained significantly more weight than women with weight gain within or below the recommendation (Fig. 1).

In the linear regression model of pregnancy weight gain (Table 4), no strong predictors of pregnancy weight retention were noted. The final model showed that, unadjusted, each 1-lb increase in pre-pregnancy BMI was associated with a decrease of 0.63 pounds of weight gain during pregnancy. However, when a correlation analysis was performed it was noted that subjects with a BMI less than 26 had a risk to gain less than 25 pounds during pregnancy that was 0.4 times that of subject with BMI more than 29. We estimate that the relative risk to gain less than 25 pounds lies between 0.2 and 0.7 with 95% confidence. This result is statistically significant ($P = 0.01$) or (RR = 0.4, CI = 0.2−0.7, $P = 0.01$). In conclusion, subject with pre-pregnancy BMI < 26 are more likely to gain more pounds during pregnancy compared to those with pre BMI > 29.

In the linear regression model (Table 5), pre-pregnancy BMI and gestational weight gain were the strongest predictors of postpartum weight retention at one year postpartum. The final adjusted model showed that each unit increase in pre-pregnancy BMI was associated with an increase of 1.23 pounds of weight retention at one year postpartum. When adjusted for age, smoking, pregnancy complication and contraception use at six months postpartum, adolescent pregnancy weight gain was a significant predictor of weight retention at 1 year postpartum. Every pound increase in pregnancy weight gain was associated with an increase of 0.37 pounds of postpartum weight retention at one year.

Discussion

Summary of Results

Pre-pregnancy BMI and gestational weight gain are predictors of postpartum weight retention in

Table 3. Number and Percent of Sample below, at and above IOM Weight Gain Recommendations

Initial Body Mass Index (BMI)	Percent of Sample (n)	Below IOM Recommendation %	At IOM Recommendation %	Above IOM Recommendation %
<19.8 (low)	19 (20%)	31.6	42.1	26.5
19.8−26.0 (normal)	52 (54%)	34.6	28.5	36.5
26.1−29.0 (high)	18 (9%)	16.7	16,7	66.5
>29 (obese)	8 (8%)	12.5	62.5	25
Total	97 (100%)*	27	30	36

*N (5) missing data.

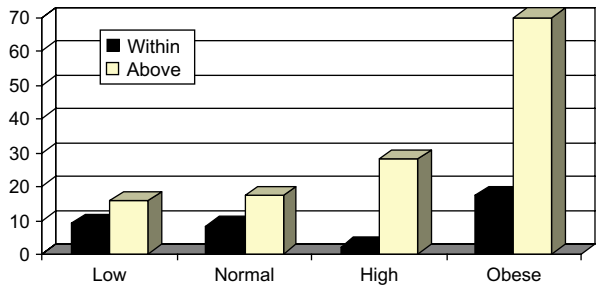


Fig. 1. Mean Weight Retention by Pre-Pregnancy BMI and Pregnancy Weight Gain According to IOM Weight Gain Recommendations. X axis, Pre-Pregnancy BMI and IOM classification; Y axis, weight gain (pound).

adolescents. Our study of 102 pregnant adolescents found that those with higher pre-pregnancy weight or those whose gestational weight gain was above the IOM recommendation retained significantly more weight than adolescent women with weight gain within or below the recommendation. These findings were independent in age, post partum contraceptive use, and smoking status.

Four in 10 adolescent girls get pregnant at least once before the age of 20 in the United States, leading to nearly 900,000 teenage pregnancies each year.[1–3]

Most adolescents who become pregnant are still growing. Normal adolescent growth is associated with increased weight gain and fat storage.[8] Compared to adolescent boys, adolescent girls have a greater proportion of body fat, with adipose tissue distributed along the upper arms, thigh, and upper back. In adolescent girls younger than 16 years of age, the annual increase in BMI is driven primarily by changes in fat-free mass; however, after 16 years of age, the increase in BMI in girls is largely because of increase in fat mass.[8] The onset of puberty is controlled by many factors that remain incompletely understood.[9] Leptin has been proposed as the hormone responsible for the initiation and progression of puberty. Leptin is produced largely in adipocytes; large fat cells produce more leptin than do small ones, and serum leptin concentrations are highly correlated with body fat content. Higher serum leptin concentration in girls is associated with increased body fat and earlier onset of puberty.[10] In a study, a 1ng/ml increase in serum leptin lowered the age at menarche by 1 month, and a gain in body fat of 1 kg lowered the timing of menarche by 13 days.[11] When menarche occurs before the individual reaches the age of 12 years, it is associated with increases in weight and body mass index.[12,13] Hence, girls who mature early are at an increased risk for obesity when compared with late maturers. Hence, early mature girls (menarche <12 yr) who get pregnant are even at a higher risk of weight gain given the already stated facts above. Scholl and colleagues have shown that the amount of postpartum weight retained after one year is significantly greater in still-growing gravidas than in other pregnant women.[14] Research comparing maternal weight gain differences among adolescents and other women has shown that adolescents gain more weight during pregnancy than their older counterparts.[15–18] Increased maternal weight gain in adolescents compared with older women has been attributed, in part, to the adolescent body working to meet the needs of its own growth in addition to the needs of the growing fetus.[18–20] As a result, some of the increased maternal weight gain in adolescents may be associated with increased birth weight of infants and more favorable pregnancy outcomes.[21–23] However, there is evidence that adolescent mothers with excessive maternal weight gain have a greater tendency to retain excess weight postpartum.[4]

Recommendations for gestational weight gain have increased in recent decades. In 1970, the National Academy of Sciences recommended a gestational weight gain of 20–25 pounds.[24] In 1983, the first edition of *Guidelines for Perinatal Care* called for a maternal weight gain of 22–27 pounds.[25] In a report issued in 1990, the IOM recommended a maternal weight gain of 25–35 pounds for women with normal weight for height[26] (Table 1). In terms of the IOM recommended weight gain, using age appropriate

Table 4. Longitudinal regression model for pregnancy weight gain

Predictors	Crude: Change in weight (lb) during pregnancy	Crude: *P* value	Adjusted: Change in weight (lb) during pregnancy	Adjusted: *P* value
Age*	0.07	0.93	0.48	0.66
Pre BMI	−0.63	0.06	−0.56	0.10
Smoking	2.67	0.37	2.43	0.46
Menarche*	−1.26	0.16	−1.04	0.28
Complications	5.57	0.10	5.5	0.13

*Change per year increase, Change per Unit and BMI.

Table 5. Linear Regression Model for Weight Retention

Predictors (for each unit increase)	Unadjusted: Change in weight retention (Lb) postpartum	Unadjusted: *P* value	Adjusted: Change in weight retention (Lb) postpartum	Adjusted: *P* value
Age*	−1.39	0.30	−2.48	0.06
Pre BMI	1.23	0.005	1.33	0.002
Smoking	−1.88	0.63	−5.41	0.19
Complications	−2.23	0.61	−5.51	0.16
Contraception	−10.34	0.02	−6.95	0.11
Weight gain	0.37	0.004	0.47	0.0002

*Change per year increase, Change per Unit and BMI.

ranges might not alter the recommended weight gain during the pregnancy. Based on the BMI for Caucasian and African American children (6−18 y), adolescent girls between the ages 15−18 with BMI ranging from 19.6 to 20.8 respectively will fall below the 50th percentile and will be considered to have normal BMI.[27] Adolescent girls of the same age group with BMI ranging 26.0−27.9 will be at >85% for BMI and be considered overweight; whereas same aged girls with BMI 27.0−29.3 will be > 95% and be considered obese. For the age group (primarily 15−18 yrs) in this study (Table 1), the ''normal'' BMI value for adolescents is also considered normal by the Institute; the range considered ''overweight'' in adolescents falls in the ''high'' category to the IOM. For this age range in adolescents ''obesity'' is a BMI ≥ 27 whereas for the IOM ''obesity'' falls in the BMI range of >29. However, the recommended pregnancy weight gain to the IOM for obese women is at least 15 pounds which is the minimum amount recommended for girls with in the ''high'' (BMI 26.1−29) to the IOM. In conclusion, there exists little difference in the BMI category (normal, overweight, obese) used primary care pediatrics compare to the IOM BMI category (normal, high, obese) respectively. Hence, the recommended weight gain during pregnancy might not alter using age appropriate ranges.

The trend toward higher recommendations for weight gain during pregnancy is motivated by concerns about the weight and health of the infant,[26] although apprehension about retaining weight after the baby is born may make the prospect of gaining more weight unattractive to some women.

Excessive maternal weight gain has been associated with many pregnancy-related health risks, including labor and delivery complications, maternal anemia, preterm labor, macrosomia, and infant mortality.[21,26,28] Proposed maternal weight gain by IOM is geared more toward adult women. Whether the recommended 25−35 pound pregnancy weight gain is adequate or too much for pregnant adolescent is undetermined and remains a topic of discussion. A larger sample of pregnant adolescents is needed to support or refute these IOM recommendations for this population. Further, post-partum weight retention, a direct consequence of excessive maternal weight gain, can lead to obesity, which increases a woman's risk of major health problems such as hypertension, heart disease, diabetes, and some types of cancer.[15,18,29] Indeed, obesity is increasing rapidly among women all over the world[30] and more women of reproductive age are becoming overweight and obese.

Obesity is steadily increasing in women of reproductive age[31] and has become a major public health concern in the United States. Reducing the proportion of children, adolescents, and adults who are obese has been listed as one of the Healthy People 2010 objectives. According to the National Health and Nutrition Examination Survey (NHANES) which has monitored changes in body weight among children and adolescents in the US since 1960, 16% of adolescents are now obese. This figure represents a threefold increase in prevalence rates since NHANES I (1976).[32]

In a study examining the potential implications of compliance with the IOM guidelines for weight gain during pregnancy, for postpartum weight retention, Keppel and Taffel found that weight retention following delivery increased as weight gain increased, and Black women retained more weight than White women with comparable weight gain during pregnancy.[4] The median weight retained among White women who gained the amount now being recommended was 1.6 pounds versus 7.2 pounds for Black women.

Adolescents are at high risk based on their ongoing growth and development of gaining an excessive amount of weight during pregnancy and should be monitored during pregnancy by professional dieticians. These findings, along with those of Keppel and Taffel, suggest that adolescent women in general, and more specifically, Black young women, are in need of advice of about how to lose weight following delivery.

Strength of Study

A major strength of our study is its ability to sample a group of adolescents from racial and ethnic minority populations, a group at higher risk for future obesity. A second strength of the study is the use of the medical record and measured weights as opposed to subject self- report, and associated recall bias.

Limitation

This study is limited by its small sample size and inability to generalize the result to a different population.

Implications

These findings suggest the need to be concerned about weight before girls get pregnant; interventions to prevent excessive weight gain during pregnancy must be in place during prenatal care. Pre-pregnancy adolescents who gain an excessive amount of weight during their first pregnancy may retain some of this weight. BMI and gestational weight gain are two predictors that must be monitored systematically with the aim of preventing postpartum obesity and the diseases that follow. Further research into understanding the relative contributions of age and parity on prenatal weight gain may be important for determining how to best prevent obesity among adolescent mothers.

Acknowledgments: This project is funded by NIH 1APH PA002033-01 DHHS Office of Population Affairs-OAPP (Office of Adolescent Pregnancy and Parenting).

References

1. Miller FC: Impact of adolescent pregnancy as we approach the new Millennium. J pediatr Adolescent gynecol. 2000; 13:5−8
2. Kirby D: Emerging answers: Research findings on programs to reduce teen pregnancy. Washington, DC, National Campaign to prevent teen pregnancy, 2001
3. Jonathan D: Klein and the Committee on Adolescence. Adolescent Pregnancy: Current Trends and Issues. Pediatrics 2005; 116:281−6
4. Keppel KG, Tafel SM: Pregnancy-related weight gain and retention: Implications of the 1990 Institute of Medicine guideline. Am J Public Health 1993; 83:1100−3
5. Blum PW: Adolescent Health Care-Clinical Issues, 83. New York, Academic Press, 1982
6. Stunkard A, Mendelson M: Obesity and the Body Image: I. Characteristics of Disturbances in the Body Image of Some Obese Persons. American Journal of Psychiatry, 1967 April; 123(10):1296−300
7. Institute of Medicine, Subcommittee on Nutritional Status and Weight Gain during Pregnancy, Nutrition during Pregnancy. Washington, DC, National Academy of Sciences, 1990
8. Abbassi V: Growth and normal puberty. Pediatrics 1998; 102:507
9. Grumbach MM, Sizonenko PC, Aubert ML: Control of the onset of puberty. In: Brown C, editor. Baltimore, Williams &Wilkins, 1990
10. Matkovi V, Ilich JZ, Skugor M, et al: Leptin is inversely related to age at menarche in human females. J lin Endocrinol Metab 1997; 82:3239
11. Ahima RS, Prabakaran D, Mantzoros C, et al: Role of leptin in the neuroendocrine response to fasting. Nature 1996; 382:250
12. Wellens R, Malina R, Roche A, et al: Body size and fatness in young Adults in relation to age at menarche. Am J Hum Biol 1992; 4:783
13. Kaplowitz PB, Slora EJ, Wasserman RC, et al: Earlier onset of puberty in girls: relation to increased body mass index and race. Pediatrics 2001; 108:347
14. Scholl TO, Hediger ML, Schall Jl, Khoo CS, Fisher RL: Maternal growth during pregnancy and the competition for nutrients. Am J Clin Nutr 1994; 60:183−8
15. Johnson JWC, Longmate JA, Frentzen B: Excessive maternal weight and pregnancy outcome. Am J Obstet Gynecol 1992; 167:353−72
16. Hediger ML, Scholl TO, Ances IG, Belsky DH, Salmon RW: Rate and amount of weight gain during adolescent pregnancy: Associations with maternal weight-for-height and birth weight. Am J Clin Nutr 1990; 52:793−9
17. Sohlström A, Forsum E: Changes in adipose tissue volume and distribution during reproduction in Swedish women as assessed by magnetic resonance imaging. Am J Clin Nutr 1995; 61:287−95
18. Kuczmarski RJ: Prevalence of overweight and weight gain in the United States. Am J Clin Nutr 1992; 55:495S−502S
19. Greene GW, Smiciklas-Wright H, Scholl TO, Karp RJ: Postpartum weight change: How much of the weight gained in pregnancy will be lost after deliver? Obstet Gynecol 1988; 71:701−7
20. Ohlin A, Rossner S: Maternal body weight development after pregnancy. Int J Obesity 1990; 14:159−73
21. Hytten FE: Weight gain in pregnancy. In: Hytten FE, Chamberlain G, editors. Clinical physiology in obstetric. London, Blackwell, 1980
22. Parham ES, Astrom MF, King SH: The association of pregnancy weight gain with the mother's postpartum weight. J Am Diet Assoc 1990; 90:550−4
23. Parker JD, Abrams B: Differences in postpartum weight retention between black and white mothers. Obstet Gynecol 1993; 81:768−74
24. Committee on Maternal Nutrition/Food and Nutrition Board, National Research Council. Maternal Nutrition and the Course of Pregnancy. Washington, DC, National Academy of Sciences, 1970
25. American Academy of Pediatrics, American College of Obstetricians and Gynecologist: Guidelines for Perinatal Care. Evanston, III, American Academy of Pediatrics, 1983
26. Subcommittee on Nutritional Status and Weigh Gain During Pregnancy. Nutrition during pregnancy. Washington, DC, Institute of Medicine, National Academy of Sciences, 1990
27. Hammer LD, Kraemer HC, Wilson DM, et al: Am J Dis Child 1991; 145:259
28. Lawrence M, McKillop FM, Dumin JVGA: Women who gain more fat during pregnancy may not have bigger babies: Implications for recommended weight gain during pregnancy. Br J Obstet Gynaecol 1991; 98:254−9
29. Segel JS, McAnarney ER: Adolescent pregnancy and subsequent obesity in African-American girls. J Adolesc Health 1994; 15:491−4
30. WHO Obesity: Preventing and Managing the Global Epidemic. Report of WHO Consultation on Obesity, 3−5 June. Geneva, WHO 1997.6
31. Howie DL, Parker JD, Shoendorf CK: Excessive maternal weight gain patterns in adolescents. J Am Diet Assoc. 2003; 103:1653−7
32. Hedley AA, Ogden L, Johnson CL, et al: Prevalence of overweight and obesity among US children, Adolescents, and Adults, 1999-2002. JAMA 2004; 291:2847−50

ELSEVIER

Journal of Clinical Epidemiology 61 (2008) 951–957

Journal of Clinical Epidemiology

Angina pectoris is a stronger indicator of diffuse vascular atherosclerosis than intermittent claudication: Framingham study

W.B. Kannel[a,*], J.C. Evans[b], S. Piper[d], J.M. Murabito[a,c]

[a]*NHLBI's Framingham Heart Study, Framingham, MA, USA*
[b]*Statistics and Consulting Unit, Boston University, Boston, MA 01702-5827, USA*
[c]*Section of General Internal Medicine, Boston University School of Medicine, Boston, MA, USA*
[d]*University of Houston, The Honors College, Houston, TX 77204, USA*

Accepted 12 October 2007

Abstract

Objective: To compare implications of Angina Pectoris (AP) and Intermittent Claudication (IC) as indicators of clinical atherosclerosis in other vascular territories.

Study Design and Setting: Prospective cohort study of cardiovascular disease (CVD) in 5,209 men and women of Framingham, MA, aged 28–62 years at enrollment in 1948–1951, who received biennial examinations during the first 36 years of follow-up. Comparative 10-year incidence of subsequent atherosclerotic CVD in participants with IC and AP relative to a reference sample free of CVD was determined.

Results: On follow-up, 95 CVD events occurred in 186 participants with IC and 206 of 413 with AP. After age, sex, and risk-factor adjustment, the proportion acquiring other CVD was 34.0% for IC and 43.4% for AP. Relative to the reference sample, those with IC had a 2.73-fold higher age and sex-adjusted 10-year hazard of CVD (95% CI 2.21, 3.38) and for AP was 3.17 (95% CI 2.73, 3.69). CVD hazard ratios remained more elevated for AP and statistically significant after standard risk factor adjustment. Risk factors accounted for more of the excess CVD risk associated with IC (34.8%) than AP (9.5%).

Conclusion: AP is as useful as IC as a hallmark of diffuse atherosclerotic CVD and an indication for comprehensive preventive measures.

Keywords: Angina pectoris; Intermittent claudication; Prognosis; Cohort study; Cardiovascular disease; Vascular damage indicator

1. Introduction

Peripheral artery disease, manifested as intermittent claudication (IC) or abnormal ankle–brachial index is an accepted marker of diffuse atherosclerotic vascular disease and increased risk for mortality, primarily from cardiovascular causes [1–9]. Angina pectoris (AP), another transient ischemic condition provoked by exertion, is regarded chiefly as a hallmark of impending myocardial infarction or a coronary fatality. For example, the Rose angina questionnaire, devised for epidemiological investigation of angina, was tested chiefly as a predictor of coronary morbidity and mortality [10–12].

We had the opportunity to examine the prognostic implications of these transient ischemic conditions, using population-based data derived from the Framingham Study between 1949 and 1990, a period during which there were few effective cardiovascular disease (CVD) therapies available or in widespread use, allowing unbiased estimates. This report compares the total atherosclerotic cardiovascular outlook of participants experiencing an initial IC or AP event, during the first 36 years of the study with a reference group of participants, free of CVD, drawn within the same calendar time frame.

2. Methods

2.1. Study sample

The Framingham Study is an ongoing, prospective cohort study of the epidemiology of CVD. From 1948 to 1953, 5,209 men and women, between the ages of 28 and 62 years, residing in Framingham, MA were enrolled and have been re-examined biennially since the study inception.

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. This work was supported by the National Heart, Lung and Blood Institute's Framingham Heart Study (Contract No. N01-HC-25195).

* Corresponding author. 73 Mount Wayte Avenue, Suite 2, Framingham, MA, USA.

E-mail address: billkannel@yahoo.com (W.B. Kannel).

0895-4356/08/$ – see front matter
doi: 10.1016/j.jclinepi.2007.10.025

What is new?

Angina pectoris joins intermittent claudication as a robust indicator of diffuse atherosclerotic vascular involvement. Both require targeting of the underlying accelerated atherogenesis in their management in addition to protecting the limbs for claudication and the heart for angina.

Details of the original sampling have been published previously [13]. Participants with no prior history of CVD at entry were eligible for the present study. First CVD events classified as either IC or AP, occurring prior to January 1, 1980, and with no other CVD event on the same day, comprised the two exposure groups of interest. When the exact date of onset of IC or AP was unknown, the mid-point between the examination of diagnosis and the last attended examination free of symptoms was used. Participants were followed for up to 10 years for new CVD events. *Participants who had not attended an examination within 2 years prior to IC or AP diagnosis were excluded from analysis ($n = 46$) because covariate data was taken from the examination prior to diagnosis.*

The sampling scheme used to select the reference group was designed such that the distribution of age, sex, and calendar time would resemble that of the IC and AP groups. For each participant in the study, one examination was chosen, at random, from all examinations attended prior to 1990 free of CVD, to serve as the baseline for the 10-year follow-up period. *If, after 10-years, the participant attended another exam, still free of CVD, a new baseline was established along with a corresponding 10-year follow-up period. This technique is the time dependent Cox regression* (SAS Institute Inc. SAS/STAT User's Guide, Version 8. Cary, NC: SAS Publishing; 1999, 2569–2656).

2.2. Diagnostic criteria for AP and IC

The clinical criteria used for the diagnosis of cardiovascular events in the Framingham Heart Study have been reported in detail [14]. A physician-administered structured questionnaire was used to determine the presence of AP and IC based on the participants' subjective manifestations. Brief, recurrent chest discomfort of up to 15 min in duration, brought on by exertion and relieved by rest or nitroglycerin, was accepted as angina if two physicians agreed that it was present. Abnormality of the ECG, either at rest or on exercise testing, was not taken into account. Participants with angina who exhibited no other manifestations of coronary heart disease (CHD) at the current examination or any prior examination were designated as having uncomplicated AP.

The presence of IC was accepted if the interviewing physicians agreed that the participant had experienced a transient, short duration, cramping discomfort in the calf that was provoked by walking, appeared sooner when the participant walked faster or uphill, and was promptly relieved by rest [14]. Abnormalities on noninvasive testing such as ankle–brachial blood pressure testing or ultrasonography were not taken into account in the evaluation of IC.

2.3. Ascertainment of CVD outcomes and death

Participants were followed for 10 years after the onset of IC or AP for the occurrence of a cardiovascular event or death. Information regarding events was obtained from biennial examinations and hospital surveillance. CVD included CHD defined as AP, coronary insufficiency (prolonged chest pain in the presence of reversible electrocardiographic changes), myocardial infarction, or CHD death; stroke or transient ischemic attack; and congestive heart failure. CVD events were adjudicated by a panel of three senior investigators (or a panel of study neurologists for cerebrovascular disease events) using standardized criteria previously reported [15]. All deaths were reviewed and cause of death was classified by the endpoint panel as due to CHD, stroke, other CVD, cancer, other causes, or unknown cause.

2.4. Covariate measures

At each biennial examination, resting blood pressures were taken with a mercury sphygmomanometer and a 14-cm cuff on the left arm of participants and readings were recorded to the nearest even number. For the purposes of this investigation, hypertension was designated if the mean of two physician obtained blood pressures was $\geqslant$140/90 mm Hg or the participant reported taking antihypertensive medications. Serum cholesterol after exam 2 in 1952 was determined by the Abell–Kendall method. Blood glucose was measured on a casual specimen of whole blood using the Nelson method. Diabetes mellitus was diagnosed if there was a casual blood glucose level of $\geqslant$200 mg/dl or the use of insulin or oral hypoglycemic medication. Body mass index was calculated from height and weight measurements obtained from trained technicians using the formula: weight(kg)/height(m^2). Current cigarette smoking was self-reported. Participants who reported regular cigarette smoking within the year prior to the examination were classified as current smokers.

2.5. Statistical analysis

The IC, AP, and reference groups were followed for up to 10 years from baseline for the occurrence of new CVD events. Endpoints of interest included the first new event, fatal or nonfatal, in each of the three types of CVD examined, CHD, congestive heart failure (CHF), and cerebral vascular accident (CVA) or transient ischemic attack, as well as the first new CVD event overall. *Events in subjects in the IC group included AP while events in the AP group*

included IC. Subject characteristics at baseline for the IC, AP, and reference group were compared after adjustment for age and sex. For continuous variables, linear regression was used, least squares mean values were produced, and the partial *F*-statistic was used to test for differences between the groups. Dichotomous variables were standardized using the direct method and the Pearson chi-square was used to test for differences between the groups.

To assess differences in CVD incidence between the IC group, the AP group, and the reference group, Kaplan–Meier survival methods were used to construct incidence (1-survival) curves (Fig. 1). As indicated in Figure 1, the age- and sex-adjusted incidence of angina pectoris is greater than that for intermittent claudication throughout the follow-up. The Group Prognostic Method [16] was used to produce group-specific, predicted incidence rates adjusting for age and sex and with further adjustment for systolic blood pressure, hypertension treatment, body mass index, current cigarette smoking, diabetes, and serum total cholesterol. Cox proportional hazards regression was used to compare CVD risk in the IC group and the AP group, both relative to the reference group. Separate models were performed for men and women and then pooled.

3. Results

Table 1 displays the baseline characteristics of the study sample. The IC and AP subjects were older and more male than the reference sample free of CVD. In addition, as expected, almost all risk factor levels were significantly ($P < 0.001$) higher in those with IC or AP than in the reference sample. Comparing IC subjects with those having AP, cigarette smoking prevalence was lowest in subjects with angina and BMI was lowest in IC subjects. Cigarette smoking and diabetes prevalence were substantially higher in the IC group than in the AP group. Those with AP were heavier than those with IC whose BMI was lower than that of the reference sample.

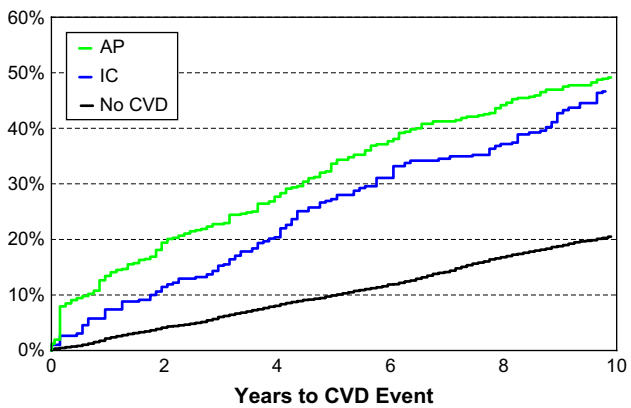


Fig. 1. Cardiovascular disease incidence in the Intermittent Claudication, Angina Pectoris, and Reference Group: Results of Age- and Sex-adjusted Kaplan–Meier Survival Analysis.

The number of men and women in the population at risk with AP was substantially larger than for IC. As a consequence, the number of other cardiovascular events over the 10-year period of follow-up was substantially larger for AP (206) than IC (95). However, the age and sex-adjusted proportion developing CVD events was about 46% compared to 20.5% in the reference group (Table 2). After multivariable risk-factor adjustment there was a greater proportion of AP victims (43.4%) having an event than observed for persons with IC (34.0%) (Table 2). This appears to be attributable to the fact that a greater fraction of the IC events (34.8%) can be ascribed to accompanying cardiovascular risk factors than for AP (9.5%).

CHD was the most common cardiovascular hazard for both AP and IC, which adjusting for age and sex, occurred at the same 35% 10-year rate. However, as expected, on adjustment for accompanying CVD risk factors the coronary rate for AP was somewhat greater (30.8%) than for IC (25.2%). Age and sex-adjusted stroke rates appeared to be almost twice as high for IC as for AP. However, on adjustment for accompanying CVD risk factors, the 10-year stroke rates were more comparable (14.5% vs. 11.2%) but the stroke rate for AP was identical to that for the reference sample (11.1%). Contrary to expectation, the 10-year heart failure rate for IC (10.8%) was similar to that for AP (12.4%) on adjustment for accompanying CVD risk factors.

Table 2 also indicates that subjects with either one of these transient exercise-induced ischemic conditions are at about four to fivefold increased risk of developing the other. The absolute age and sex-adjusted risk of IC in subjects with AP (16.3%) is greater than the 9.8% risk of AP in subjects with IC. This greater likelihood of angina patients to develop IC than the converse persists after adjusting for associated risk factors.

Table 3 displays the age and sex-adjusted and multivariable risk-factor adjusted hazard ratios for other cardiovascular events in subjects with IC and AP. For IC, the age and sex-adjusted hazard of developing angina was greatest (4.3-fold) followed by heart failure (3.1-fold), and then coronary disease (2.7-fold) and least of all stroke (2.1-fold). This ranking order persists after multivariable risk factor adjustment. Overall, IC imposes a 2.7-fold age and sex-adjusted increased risk of other cardiovascular events, which is reduced to 1.8-fold after multivariable risk-factor adjustment.

For AP, the age and sex-adjusted hazard ranking is greatest for IC (5.2-fold) followed by other manifestations of coronary disease (3.1-fold), heart failure (2.7-fold), and least stroke (1.1-fold). This rank order holds after multivariable risk-factor adjustment. Overall, AP carries a 3.2-fold age and sex-adjusted hazard of other cardiovascular events, which is reduced to 2.9-fold after multivariable risk-factor adjustment.

Analysis of the multivariable hazard provides insight into the amount of influence that the associated CVD risk factors have in generating the increased propensity to other

Table 1
Baseline characteristics according to first CVD event as intermittent claudication or angina pectoris: Framingham Heart Study, 1948–1979

Characteristic	Intermittent Claudication; $n = 186$	Angina Pectoris; $n = 413$	Reference Sample Free of CVD; $n = 6,860$
Age (years)	62.5 (8.9)	60.8 (8.6)	56.8 (10.8)
Men (%)	56	48	43
Systolic blood pressure (mm Hg)[a]	146.1 (1.60)	144.6 (1.07)	136.9 (0.27)
Diastolic blood pressure (mm Hg)[a]	83.1 (0.87)	86.6 (0.59)	82.4 (0.14)
Hypertension treatment[a]	25.1	17.5	11.1
Total cholesterol (mg/dL)[a]	248 (3.3)	252 (2.2)	235 (0.5)
Diabetes mellitus[a]	16.7	4.5	3.4
Cigarette smoking, current[a]	60.0	37.6	40.7
Body mass index (kg/m^2)[a]	25.4 (0.31)	27.2 (0.21)	26.0 (0.05)

All *P*-values for difference between the groups ≤0.001.
[a] Age and sex-adjusted least squares means (± standard error) for continuous variables; direct age and sex-adjusted prevalence for categorical variables.

CVD. Judging by the size of the reduction in the hazard ratio from age-adjusted to multivariable-adjusted models, about 34.8% of the CVD potential of IC is attributable to the associated risk factors, whereas for AP only about 9.5%.

Table 4 provides sex specific hazard ratios for development of cardiovascular events in subjects presenting with AP and IC. For subjects with IC, the total CVD hazard ratios are greater for women than men for all outcomes, with the possible exception of stroke (not statistically significant). For AP subjects, the hazard ratios for CVD are comparable in men (hazard ratio [HR] 2.9) and women (HR 2.8). The hazard of IC in women with AP (HR 10.5) is substantially greater than for men (HR 2.7).

Table 2
Cardiovascular events according to intermittent claudication and angina pectoris status: Framingham Study cohort: men and women combined—10-year incidence rates

Events/person years Predicted 10-year rates	Intermittent Claudication	Angina Pectoris	Reference Sample
CVD	95/1,269	206/2,719	983/45,054
Unadjusted	53.8%	51.1%	19.9%
Age and sex	45.7%	47.2%	20.5%
Multivariable	34.0%	43.4%	21.3%
CHD	67/1,392	143/2,978	602/40,526
Unadjusted	39.9%	36.1%	13.9%
Age and sex	34.9%	34.6%	14.2%
Multivariable	25.2%	30.8%	14.9%
CVA	38/1,517	42/3,446	419/40,441
Unadjusted	22.9%	11.8%	10.2%
Age and sex	20.7%	11.6%	10.5%
Multivariable	14.5%	11.2%	11.1%
CHF	27/1,545	48/3,407	199/39,187
Unadjusted	16.0%	13.0%	5.1%
Age and sex	14.6%	12.8%	5.2%
Multivariable	10.8%	12.4%	5.5%
IC		27/1,500	156/37,592
Unadjusted		16.8%	4.0%
Age and sex		16.3%	4.0%
Multivariable		14.5%	4.1%
AP	35/3,383		72/36,818
Unadjusted	9.9%		1.9%
Age and sex	9.8%		1.9%
Multivariable	9.0%		2.0%

4. Discussion

Both *AP* and *IC* are transient symptoms of ischemic vascular disease brought on by exertion and relieved by rest. Angina is usually regarded as a hazard for development of a myocardial infarction or coronary fatality, which it clearly is. It is less often considered as an indicator of diffuse atherosclerotic disease involving other vascular territories as is now the case for IC [17–19]. The data presented indicates that angina as well as IC is a hallmark of diffuse atherosclerotic vascular disease involving the circulation to the limbs as well as the heart, imparting a two to threefold excess risk compared to the reference group. It appears that AP deserves as much attention as IC as an indicator of the need for comprehensive and more aggressive preventive measures against diffuse accelerated atherogenesis.

Atherothrombotic vascular disease is usually a diffuse condition involving the arterial circulation to the heart, brain, kidney, and periphery. Most of the risk factors that predispose to involvement of one arterial bed also apply to the others; consequently, it should be expected that having one clinical manifestation of atherosclerosis, increases the risk of developing the others [14]. However, the hazard of clinical events in other vascular territories is apparently not chiefly a product of shared risk factors. Judging by the size of the reduction in the hazard ratio on adjustment for the coexistent risk factors, they only account for about 35% of the other CVD risk for IC and 9.5% of the hazard of AP. The nature of the unique effect is uncertain but it is possible that the appearance of a clinical manifestation indicates greater vulnerability to the cluster of predisposing cardiovascular risk factors. Furthermore, atherosclerosis per se begets atherothrombosis by its tendency to progress on its own.

It was previously reported from the Framingham Study that more than half of persons with IC at initial diagnosis already have coexistent atherothrombotic CVD [18,20–22]. As early as 1974, the Study also reported that the chief

Table 3
CVD subsequent to intermittent claudication vs. angina pectoris as initial CVD events compared with a reference sample free of CVD at baseline, 10-year follow-up

	Sex Pooled			
	Hazards Ratios (95% Confidence Intervals)			
	Intermittent Claudication		Angina Pectoris	
Events	Age and Sex Adjusted	Multivariate adjusted[a]	Age and Sex Adjusted	Multivariable Adjusted[a]
Cardiovascular disease	2.73 (2.21, 3.38)	1.78 (1.43, 2.22)	3.17 (2.72, 3.69)	2.87 (2.47, 3.35)
Coronary heart disease	2.72 (2.11, 3.51)	1.74 (1.33, 2.27)	3.09 (2.58, 3.71)	2.66 (2.21, 3.20)
Stroke	2.13 (1.52, 2.97)	1.36 (0.96, 1.92)	1.12 (0.81, 1.54)	1.02 (0.74, 1.40)
CHF	3.08 (2.05, 4.61)	2.15 (1.41, 3.29)	2.69 (1.96, 3.69)	2.52 (1.83, 3.48)
Intermittent claudication			5.22 (3.48, 7.83)	4.60 (3.04, 6.98)
Angina pectoris	4.27 (2.82, 6.46)	3.71 (2.40, 5.74)		

[a] Adjusted for age, systolic blood pressure, hypertension treatment, body mass index, diabetes, current smoking, and serum total cholesterol.

hazard of IC was not the loss of a limb but rather a serious cardiovascular event [23]. It was suggested that IC may be a marker for predisposition to atherothrombotic events in other vascular territories, which proved to be the case [1−3].

Vascular bruits often signify diseased arteries and it is not surprising that femoral bruits have been reported to be associated with a high (20−30%) prevalence of IC. However, these same bruits are also associated with a significantly increased prevalence of CHD and heart failure [24]. Likewise, carotid bruits, indicating vascular disease of the cerebral circulation, is not only associated with a two to threefold increased stroke risk, but also carries a two to threefold increased risk of coronary disease, IC, and heart failure. Because the peripheral vessels are more accessible to noninvasive testing for obstruction to flow, there is merit in detecting presymptomatic arterial disease, so that timely preventive measures can be implemented to protect against lethal clinical manifestations of atherothrombotic disease.

The Framingham Study has crafted multivariable risk profiles for identifying high-risk persons for development of IC and coronary disease that can be used to educate patients about modifiable risk factors for avoiding CVD [22,25]. Although this has not been done specifically for AP, the coronary risk profile can be used to identify high-risk candidates for AP. Assessment of multivariable risk of AP would indicate the need for more aggressive preventive measures to prevent ischemic vascular events in other vascular territories as well those involving the heart.

It is estimated, using Framingham Study and National Heart, Lung and Blood Institute data that the prevalence of AP in the year 2001 was 6,800,000 cases. In the national data, as in the Framingham Study, age-adjusted prevalence of angina was greater in women than men. National data also indicate that the prevalence of AP is higher in blacks and Mexican Americans [26]. Peripheral artery disease affects more than 5 million persons in the United States and is also higher in racial and ethnic minorities [4,27].

IC appears to be the most underdiagnosed and least aggressively managed clinical atherosclerotic condition [28,29]. The peripheral artery disease coalition and the NHLBI are launching a campaign to remedy this public health problem (www.padcoalition.org). The major problem the patient with IC faces is the high risk of other cardiovascular events; attention to the ischemic limb is not enough. The same applies for AP, where alleviation of myocardial ischemia is not enough. Investigation of IC in

Table 4
Sex-specific CVD hazard following angina vs. intermittent claudication

	Hazards Ratios (95% Confidence Intervals)			
	Men		Women	
	Age Adjusted	Multivariable Adjusted[a]	Age Adjusted	Multivariable Adjusted[a]
Intermittent claudication				
Cardiovascular disease	2.53 (1.92, 3.30)	1.67 (1.26, 2.22)	3.32 (2.36, 4.67)	2.09 (1.47, 2.99)
Coronary heart disease	2.38 (1.73, 3.27)	1.63 (1.17, 2.28)	3.76 (2.47, 5.74)	2.14 (1.37, 3.34)
Angina pectoris	3.92 (2.32, 6.63)	3.95 (2.27, 6.89)	5.27 (2.70, 10.30)	3.72 (1.81, 7.64)
Angina pectoris				
Cardiovascular disease	3.14 (2.56, 3.84)	2.93 (2.38, 3.59)	3.24 (2.58, 4.06)	2.81 (2.23, 3.54)
Coronary heart disease	3.12 (2.48, 3.92)	2.81 (2.23, 3.55)	3.08 (2.27, 4.18)	2.39 (1.75, 3.28)
Intermittent claudication	3.25 (1.88, 5.63)	2.74 (1.57, 4.79)	11.29 (5.94,21.43)	10.49 (5.36, 20.50)

Events compared with a reference sample free of CVD at baseline, 10-year follow-up.

[a] Adjusted for age, systolic blood pressure, hypertension treatment, body mass index, diabetes, current smoking, serum total cholesterol. Sex specific hazard ratios for stroke and heart failure (not shown) were not statistically significant compared to reference sample.

the Framingham Study has identified a number of risk factors that lead to its occurrence. These include age, sex, serum cholesterol, hypertension, cigarette smoking, diabetes, and presence of CHD [14]. They are also the standard risk factors for coronary disease and stroke. For AP, systolic blood pressure, serum cholesterol, diabetes, cigarette smoking, and hematocrit are significant independent risk factors in one or both sexes [14,30]. There is evidence supporting the efficacy of risk factor correction for reducing risk of other CVD events in patients with AP and IC [31−36].

Although lacking in sensitivity and specificity, a detailed history and physical examination are important in detecting and evaluating peripheral artery disease. The ankle−brachial index can be done to confirm the diagnosis and help stratify the risk since it correlates well with disease severity, functional symptoms, and disease progression. The ankle−brachial index also predicts CVD and stroke mortality, the greatest hazard of the condition [5−7]. Transient ischemic vascular episodes involving the circulation of the heart and limbs are often silent and a significantly compromised circulation may exist without symptoms. Noninvasive testing in a population indicates that the true prevalence of peripheral artery disease is at least five times greater than would be estimated from the reported prevalence of IC [37]. Likewise, three of every four transient ischemic cardiac episodes detected by ECG monitoring in angina are silent [38−40].

5. Limitations

The Framingham Study cohort has few blacks and other minority population subgroups, limiting generalizability of the data. The number of events in the CVD subgroups of myocardial infarction, stroke, and heart failure were too few to provide confident estimates of differences of the size of hazard ratios for AP in comparison to IC in the two sexes. The use of only a clinical assessment of transient myocardial and peripheral artery ischemia allows many patients with silent occlusive arterial disease who are also at high risk of other atherosclerotic events to go undetected. Ankle−brachial index is a feasible office procedure that could be used on asymptomatic patients with an unfavorable multivariable risk profile to detect occult peripheral artery disease needing comprehensive preventive measures; especially for those aged 50−69 with diabetes and smoking or age 70 or over [41]. Comparable simple office procedures for detecting occult myocardial ischemia in high-risk persons are currently unavailable.

6. Conclusions

Once detected, cardiovascular risk factor modification, symptomatic relief, and use of antiplatelet agents form the core of the management of both angina and IC. The major cardiovascular risk factors adversely affect all vascular territories, increasing vulnerability to multiple clinical manifestation of atherosclerosis including CHD and IC [14]. Modification of risk factors intended to prevent a particular atherosclerotic cardiovascular event should also prevent other outcomes. Optimal appraisal of the hazard, aggressiveness, and urgency for treatment is best obtained from a multivariable cardiovascular risk profile that estimates the probability of a cardiovascular event given the existing constellation of predisposing factors.

References

[1] Bowlin SJ, Medalie JH, Flocke SA, Zyzanski SJ, Yaari S, Goldbourt U. Intermittent claudication in 8343 men and 21-year specific mortality follow-up. Ann Epidemiol 1997;7:180−7.

[2] Bainton D, Sweetnam P, Baker I, Elwood P. Peripheral vascular disease: consequence for survival and association with risk factors in the Speedwell prospective heart disease study. Br Heart J. 1994;72: 128−32.

[3] Smith GD, Shipley MJ, Rose G. Intermittent claudication, heart disease risk factors, and mortality. The Whitehall Study. Circulation 1990;82:1925−31.

[4] McDermott MM, Liu K, Criqui MH, Ruth K, Goff D, Saad MF, et al. Ankle-brachial index and subclinical cardiac and carotid disease: the multi-ethnic study of atherosclerosis. Am J Epidemiol. 2005;162: 33−41.

[5] Resnick HE, Lindsay RS, McDermott MM, Devereux RB, Jones KL, Fabsitz RR, et al. Relationship of high and low ankle brachial index to all-cause and cardiovascular disease mortality: the Strong Heart Study. Circulation 2004;109:733−9.

[6] O'Hare AM, Katz R, Shlipak MG, Cushman M, Newman AB. Mortality and cardiovascular risk across the ankle-arm index spectrum: results from the Cardiovascular Health Study. Circulation 2006;113:388−93.

[7] Newman AB, Shemanski L, Manolio TA, Cushman M, Mittelman M, Polak JF, 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.

[8] Newman AB, Sutton-Tyrrell K, Vogt MT, Kuller LH. Morbidity and mortality in hypertensive adults with a low ankle/arm blood pressure index. JAMA 1993;270:487−9.

[9] Newman AB, Siscovick DS, Manolio TA, Polak J, Fried LP, Borhani NO, et al. Ankle-arm index as a marker of atherosclerosis in the Cardiovascular Health Study. Cardiovascular Heart Study (CHS) Collaborative Research Group. Circulation 1993;88:837−45.

[10] Lampe FC, Whincup PH, Wannamethee SG, Ebrahim S, Walker M, Shaper AG. Chest pain on questionnaire and prediction of major ischaemic heart disease, events in men. Eur Heart J 1998;19:63−73.

[11] Bulpitt CJ, Shipley MJ, Demerovic J, Ebi-Kryston KL, Markowe HL, Rose G. Predicting death from coronary heart disease using a questionnaire. Int J Epidemiol 1990;19:899−904.

[12] LaCroix AZ, Guralnik JM, Curb JD, Wallace RB, Ostfeld AM, Hennekens CH. Chest pain and coronary heart disease mortality among older men and women in three communities. Circulation 1990;81:437−46.

[13] Dawber TR, Kannel WB, Lyell LP. An approach to longitudinal studies in the community: the Framingham Study. Ann N Y Acad Sci 1963;107:539−56.

[14] Cupples LA, D'Agostino RB, Keily D. 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. Bethesda, MD: US Dept. of Health and Human Services. Public Health Service, National Institutes of Health; 1987. NIH Publication No. 87-2703.

[15] Abbott RD. The Framingham Study: an epidemiologic investigation of cardiovascular disease, section 37: the probability of developing certain cardiovascular diseases in eight years at specified values of some characteristics. Bethesda, MD: National Heart, Lung and Blood Institute; 1987. 2004.

[16] Chiang IM, Gelman R, Pagano M. Corrected group prognostic curves and summary statistics. J Chron Dis 1982;35:669−74.

[17] Mohler ER. Peripheral arterial disease: identification and implications. Arch Intern Med 2003;163:2306−14.

[18] O'Riordain DS, O'Donnell JA. Realistic expectations for the patient with intermittent claudication. Br J Surg 1991;78:861−3.

[19] Hagman M, Wilhelmsen L, Pennert K, Wedel H. Factors of importance for prognosis in men with angina pectoris derived from a random population sample. The Multifactor Primary Prevention Trial, Gothenburg, Sweden.

[20] Kannel WB, McGee DL. Update on some epidemiologic features of intermittent claudication. J Am Geriatr Soc 1985;33:13−8.

[21] Mohler ER 3rd. Peripheral artery disease: identification and implications. Arch Intern Med 2003;163:2306−14.

[22] Murabito JM, D'Agostino RB, Silbershatz H, Wilson PW. Intermittent claudication. A risk profile from the Framingham Study. Circulation 1997;96:44−9.

[23] Peabody CN, Kannel WB, McNamara PM. Intermittent claudication: surgical significance. Arch Surg 1974;109:693−7.

[24] Kannel WB. Epidemiologic relation of cardiovascular disease among different vascular territories. 2nd Edition. Atherothrombosis and cardiovascular disease. London: Lippincott Williams & Wilkins; 2005. Chap 98 1493−1496.

[25] Wilson PWF, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. Circulation 1998;97:1837−47.

[26] National Health and Nutrition Examination Survey NHANES III (1988-1994), CHC/NCHS. Heart Disease and Stroke Statistics-2003 Update American Heart Association.

[27] Selvin E, Erlinger TP. Prevalence of and risk factors for peripheral arterial disease in the United States: results from the National Health and Nutrition Examination Survey, 1999-2000. Circulation 2004; 110(6):738−43.

[28] McDermott MM, Mehta S, Ahn H, Greenland P. Atherosclerotic risk factors are less intensively treated in patients with peripheral arterial disease than in patients with coronary artery disease. J Gen Intern Med 1997;12(4):209−15.

[29] Hirsch AT, Criqui MH, Treat-Jacobson D, Regensteiner JG, Creager MA, Olin JW, et al. Peripheral arterial disease detection, awareness, and treatment in primary care. JAMA 2001;286: 1317−24.

[30] Hagman M, Wilhelmsen L, Pennert K, Wedel H. Factors of importance for prognosis in men with angina pectoris derived from a random population sample. The multifactor primary prevention trial, Gothenberg Sweden. Am J Cardiol 1988;61:530−5.

[31] Leng GC, Price JF, Jepson RG. Lipid-lowering for lower limb atherosclerosis. Cochrane Database Syst Rev 2000; CD000123.

[32] Mehler PS, Coll JR, Estacio R, Esler A, Schrier RW, Hiatt WR. Intensive blood pressure control reduces risk of cardiovascular events in patents with peripheral artery disease and type 2 diabetes. Circulation 2003;107:753−6.

[33] Ostergren J, Sleight P, Dagenais G, Danisa K, Bosch J, Davies R, et al. Impact of rampiril in patients with evidence of clinical or subclinical peripheral arterial disease. Eur Heart J 2004;25:17−24.

[34] Yusuf S, Sleight P, Pogue J, Bosch J, Davies R, Dagenais G. Effects of an angiotensin-converting-enzyme inhibitor, rampiril on cardiovascular events in high-risk patients. The Heart Outcomes Prevention Evaluation Study Investigators. N Engl J Med 2000;342: 145−53.

[35] Pedersen TR, Kjekshus J, Pyorala K, Olsson AG, Cook TJ, Musliner TA, et al. Effect of Simvastatin on ischemic signs and symptoms in the Scandinavian simvastatin survival study (4S). Am J Cardiol 1998;81:333−5.

[36] Hirsch AT, Haskal ZJ, Hertzer NR, Bakal CW, Creager MA, Halperin JL, et al. ACC/AHA 2005 Guidelines for the management of patients with peripheral arterial disease (lower extremity, renal, mesenteric, and abdominal aortic): executive summary 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. J Am Coll Cardiol 2006;47:1239−312.

[37] Vogt MT, Cauley JA, Kuller LH, Hulley SB. Prevalence and correlates of lower extremity arterial disease in elderly women. Am J Epidemiol 1993;137:559−68.

[38] Pepine CJ. Clinical aspects of silent myocardial ischemia in patients with angina and other forms of coronary heart disease. Am J Med 1986;80:24−34.

[39] Deanfield JE, Maseri A, Selwyn AP, Ribiero P, Chierchia S, Krickler S, et al. Myocardial ischemia during daily life in patients with stable angina: its relation to symptoms and heart rate changes. Lancet 1983;2:753−8.

[40] Cecchi AC, Dovellini EV, Marchi F, Pucci P, Santoro GM, Fazzini PF. Silent myocardial ischemia during ambulatory electrocardiographic monitoring in patients with effort angina. J Am Coll Cardiol 1983;1:934−9.

[41] Murabito JM, Evans JC, Larson MG, Nieto K, Levy D, Wilson PWF. The ankle-brachial index in the elderly and risk of stroke, coronary disease, and death—the Framingham study. Archives of Internal Medicine 2003;163:1939−43.

Using the Case Mix of Pressure Ulcer Healing to Evaluate Nursing Home Performance

Alok Kapoor, MD, MSc,
Boris Kader, PhD,
Howard Cabral, PhD, MPH,
Arlene S. Ash, PhD,
Dan Berlowitz, MD, MPH

Pressure ulcer healing is an important quality measure for nursing homes, but the factors that predict healing have not been well studied. Using the Minimum Data Set, the authors identified candidate variables for a logistic regression, risk-adjustment model to predict ulcer healing. The authors then assessed model discrimination and calibration. Finally, the authors compared unadjusted with risk-adjusted performance for the individual facilities within a nursing home chain. Significant predictors of healing included mobility in bed, presence of a stage 2 ulcer (compared with a stage 4 ulcer), absence of paraplegia and quadriplegia, and absence of end-stage illness. The model C statistic was 0.67, and the calibration was acceptable. Judgments about nursing performance varied in 2 cases depending upon whether unadjusted or risk-adjusted performance was used. The model that the authors developed contains credible predictors of healing. Pressure ulcer healing may be one of many indicators used to evaluate nursing home quality. (Am J Med Qual 2008; 23:342-349)

Keywords: risk adjustment; pressure ulcer healing; nursing home

AUTHORS' NOTE: This work was presented at the following conferences: Deriving and Evaluating a Risk-Adjustment Model for Pressure Ulcer Healing that uses the Minimum Data Set Paper presented (oral presentation) at: National Research Service Award (NRSA) annual conference; June 2006; Seattle, Washington.Predicting Pressure Ulcer Healing Using the Minimum Data Set. Poster presented at: Society for General Internal Medicine (SGIM) annual national conference; April 2006; Los Angeles, CA. Predicting Pressure Ulcer Healing Using the Minimum Data Set. Poster presentation at: New England Regional SGIM annual conference; March 2006; Boston, MA.

Dr Kapoor is from the Hospital Medicine Unit, Section of General Internal Medicine, Boston University, Boston, Massachusetts. Dr Berlowitz is from the Section of General Internal Medicine, Boston University, Boston, Massachusetts and Center for Health Quality, Outcomes, and Economic Research, Edith Nourse Rogers Memorial VA Hospital, Bedford, Massachusetts. Dr Kader is from the Center for Health Quality, Outcomes, and Economic Research, Edith Nourse Rogers Memorial VA Hospital, Bedford, Massachusetts. Dr Cabral is from the Department of Biostatistics, Boston University, Boston, Massachusetts, and Dr Ash is from the Health Care Research Unit, Section of General Internal Medicine, Boston University, Boston, Massachusetts. *Corresponding author:* Dr Alok Kapoor, 801 Massachusetts Avenue, Second Floor, Division of General Internal Medicine, Boston, MA 02118 (e-mail: alok.kapoor@BMC.org).

American Journal of Medical Quality, Vol. 23, No. 5, Sep/Oct 2008
DOI: 10.1177/1062860608316109

INTRODUCTION

Assessing quality in nursing homes is becoming increasingly important. Quality can be assessed in different ways, but there has been a push to develop outcome-based quality indicators. One possible outcome is pressure ulcer healing. Pressure ulcers are areas of tissue necrosis that develop when soft tissue is compressed for a prolonged period of time. Staging is usually rated on a scale of 1 to 4, where 1 represents erythema with intact skin, and stages 2 through 4 represent increasing levels of tissue breakdown. Consistent with notions of what constitutes a good outcome to assess, pressure ulcers are common and usually treatable with good quality care.[1] To be

accepted by nursing homes being evaluated, models built to compare provider performance (ie, risk-adjustment models) must adjust the rate of pressure ulcer healing for differences in case mix. Case mix for pressure ulcer healing may include predictors related to the size and stage of the ulcer, as well as demographic and comorbid illness information. Any variation in the rate of healing not explained by case mix can be attributed to quality (or random error).[1]

Little has been published about the predictors that are important to adjust for when evaluating pressure ulcer healing. Previous investigations were limited in the number of available clinical predictors.[2] The Minimum Data Set (MDS) contains detailed information describing the clinical, behavioral, and social status of each resident. Data in the MDS are collected by registered nurses on admission, every 90 days thereafter, and for significant health events. The objectives of this study were (1) to develop a risk-adjustment model that takes advantage of the detailed, longitudinal clinical information of MDS to predict which individuals with a pressure ulcer will heal and (2) to apply this model to a nursing home chain and compare the performance of individual facilities using unadjusted versus risk-adjusted performance measures.

METHODS

Study Setting and Sample

Pressure ulcer healing was evaluated among individuals with a stage 2 or higher pressure ulcer residing in a nursing home chain between 1997 and 1998. We had information about 110 facilities within the chain, most of which are located in the southeastern part of the United States. This cohort represented the incident cases of pressure ulcer disease captured in a previous study of the predictors of pressure ulcer development.[3,4]

To be eligible for the study, residents needed at least 2 assessments (ie, 2 recordings of MDS data) available that evaluated for the presence of a pressure ulcer. The index assessment was the first assessment of each quarter of the calendar year. The outcome assessment was the assessment that was nearest to 90 days after the index assessment, separated in time by at least 45 days and up to 135 days without a discharge or transfer out of a member facility. The outcome assessment for 1 quarter could therefore serve as the index assessment for the next quarter. Residents with an outcome assessment or any interval assessment that indicated a readmission were excluded from the analyses for that quarter. An individual could contribute multiple times as a healing or nonhealing event. After identifying a sample of eligible residents, we randomly divided them into 60%:40% derivation:validation sets. Information about the chain and the collection of descriptive information about its residents have been published previously by Berlowitz et al.[4]

Defining Pressure Ulcer Healing

We defined healing as a stage 2 or higher ulcer identified on index assessment that resolved to reepithelialization on the outcome assessment.

Selection of Study Predictors

To build the risk-adjustment model, we selected the following candidate predictors from the literature on pressure ulcer healing: age, continence, quadriplegia and paraplegia, mobility, presence of fever within the last 7 days, and ulcer stage.[2,5] Because of the limited number and quality of publications about pressure ulcer healing, we also consulted the pressure ulcer development literature hypothesizing that the opposite states of predictors of pressure ulcer development would, in turn, predict healing. From this literature, we selected sex, body mass index (BMI), transfer ability, bed bound status, end-stage illness, cognitive impairment, hip fracture that has occurred within the past 180 days, diabetes, peripheral vascular disease, history of a resolved ulcer, and presence of edema.[3,6-8] From our clinical insight and the larger wound care literature, we also chose bowel continence, congestive heart failure, cancer, and human immunodeficiency virus (HIV).[9-11]

We did not include the predictors that represented a process of care, such as prescription of a pressure relieving device, rehabilitation, or ulcer dressing. The use of these processes relates to the quality of the care being administered and therefore should not be controlled for in the analysis. Similarly, race was not included in our baseline analysis given that there is no biological reason that individuals of a certain race should heal differently from another. Controlling for race could excuse facilities that cared poorly for their minority residents. As a sensitivity analysis, we examined the effect of race on ulcer healing. In contrast, the development of a pressure ulcer was predicted in male sex in a previous study and therefore was included in the current model.[3]

Analyses

We performed all analyses using Statistical Analyses System (SAS), version 9.1 (SAS Institute, Cary, North Carolina). First, we examined the bivariate association of candidate predictors with the binary outcome healing, yes or no. This information was used to determine whether to model predictors as continuous or categorical predictors and when to collapse categories (eg, if there was no difference in effect among 2 risk categories). Predictors with sufficient distribution across clinically relevant predictor levels and which had sufficient nonmissing data were chosen for multivariable regression. Selection of the predictors to include in the multivariable model was not based on the significance level of bivariate associations, in accordance with recommendations by Sun et al.[12]

Next, we entered candidate predictors into a logistic regression model. The C statistic was examined to assess model discrimination (ie, ability to differentiate high-risk and low-risk patients), and the Hosmer-Lemeshow goodness of fit statistic was used to assess model calibration (ie, how well the model-predicted risk matched the actual probability of healing). More specifically, we evaluated calibration by dividing the sample into deciles of increasing model-predicted risk, or probability, of healing. Within each decile, we compared the expected rate calculated from logistic regression with the observed rate. Because a single patient could contribute up to 6 observations, we repeated these analyses using generalized estimating equations from the GENMOD procedure in SAS. As results did not change appreciably using GENMOD, we report only the results from logistic regression.

We tested several models with variables progressively added and deleted using the above model performance statistics and our a priori beliefs about which variables were important to retain in the model. Specifically, we decided to retain age, male sex, bed mobility, urinary continence, and ulcer stage in all models because the preponderance of evidence supported their inclusion. We also included any variable that met statistical significance at the usual alpha 0.05 level. Of the remaining variables, we included any that obtained a substantial association with healing (10% increase or decrease in odds) and that we felt to represent an important predictor of healing based on our clinical insight and the medical literature. Also, among the variables that did not meet statistical significance, we excluded variables that had counterintuitive results, such as cancer, predicting more healing. Counterintuitive associations are unlikely to predict outcomes in external settings, which is an important goal when building risk-adjustment models.

We conducted a sensitivity analysis of our results, examining duration of the assessment period as a predictor of healing. We also analyzed a model with ulcer stage alone given the strength of the association for this predictor in previous analyses. Finally, we evaluated the model in the validation set. Specifically, we applied the regression coefficients from the model generated from the derivation set to the validation set. We again calculated the C statistic and Hosmer-Lemeshow χ^2 test. After model validation, we repeated the multivariable regression on the entire dataset and documented the individual variable associations calibrated to the entire data.

Next, we applied our risk-adjustment model for the evaluation of ulcer healing in individual facilities that cared for a minimum of 30 ulcer episodes. For each facility, we calculated 2 rates; the observed rate that was the average rate of healing for residents within a nursing home, and the expected rate that was the average of the probabilities predicted by the model for each resident. To test whether there was a difference in case mix among facilities, we performed an analysis of variance (ANOVA) on the expected rate of pressure ulcer healing. We also examined the model to see whether judgments about nursing home performance differed when comparing the unadjusted versus risk-adjusted measures. Specifically, we compared the observed rates of healing for each facility with its acceptability interval. The acceptability interval is the 95% confidence interval (CI) around the expected rate. Facilities with an observed rate falling above the interval represent top performing outliers, and those with rates falling below the interval represent bottom outliers.

RESULTS

In total, there were 4327 eligible ulcer episodes cared for by 110 nursing home facilities; 3931 ulcer episodes (representing 2666 subjects) had complete data available. Missing data occurred primarily in comorbid illness. We did not have information on why the data were missing in these cases except to note that the same 395 subjects did not have any comorbid illness information; therefore, we did not attempt to impute their missing values. Sixty percent of included observations represented first ulcer episodes recorded for a

Table 1

Comparison of Descriptive Statistics for Derivation and Validation Set of Ulcer Episodes

	Derivation Set, n = 2384 (%)	Validation Set, n = 1547 (%)
Age < 65 y	186 (8)	135 (9)
Male	719 (30)	478 (31)
White race	2047 (86)	1326 (86)
Stage 2 ulcer	1536 (64)	985 (64)
Stage 3 ulcer	504 (21)	324 (21)
Intact bed mobility	1268 (53)	863 (56)
Intact transferring from bed	466 (19)	315 (20)
Urinary continence	712 (30)	473 (31)
Bowel continence	737 (31)	503 (33)
Paraplegia or quadriplegia	44 (2)	33 (2)
End-stage illness	104 (4)	75 (5)
History of resolved ulcer	431 (18)	278 (18)
Pressure ulcer healed	1367 (57)	856 (55)

subject, 29% represented second ulcer episodes, and 11% represented third or higher ulcer episodes. The duration of ulcer episode (time between recordings of MDS data) was normally distributed with a mean of 88 ± 18 days. In the derivation set, there were 2384 ulcer episodes compared with 1547 in the validation set. The random 60%:40% derivation set:validation set split of the subjects successfully divided ulcer episodes equivalently across all descriptive variables (Table 1). Sixty-five percent of the ulcer episodes involved stage 2 ulcers. Patients had a large number of physical impairments. Subjects had no bed mobility in 39% of ulcer episodes. In only 30% of ulcer episodes were subjects usually continent of urine. In only 31% were subjects continent of stool (bowel continence).

Bivariate Analysis

Several candidate predictors showed association with pressure ulcer healing in bivariate analysis. We highlight some of the more notable results of this analysis below. Lower stage ulcer was substantially associated with the outcome (Table 2). Having urinary continence at all times (as compared with always being incontinent) showed a paradoxically lower rate of healing (odds ratio [OR] = 0.77; 95% CI = 0.65-0.91). When we recoded individuals continent with an indwelling catheter as always incontinent, as Berlowitz et al did previously, we found the more logical association of continence predicting healing (OR = 1.52; 95% CI = 1.28-1.81).[4] Given the lack of a definite trend toward less healing in the occasionally

Table 2

Rate of Pressure Ulcer Healing by Stage of Ulcer

Initial Ulcer Stage	Rate of Healing (95% Confidence Interval)[a] (%)
2	67 (64-69)
3	44 (40-48)
4	32 (27-36)

[a]Unadjusted analysis

and frequently incontinent groups, we collapsed the groups with any degree of continence and compared them to those always incontinent. Body mass index was divided into 8 predictor levels, but there was no trend to the association. Therefore, we modeled BMI as a binary predictor with a cutpoint of 19, which is a commonly cited threshold for being underweight and possibly malnourished. We also considered multiple other cognitive predictors, such as delirium, periods of disordered thought, and restlessness, but there was a substantial number of incomplete data for these predictors, and so we did not further analyze these items. In addition, HIV was too rare to include in the model. In total, we analyzed 20 predictors in the multivariable phase of our analysis.

Multivariable Analysis

Among the predictors entered into the multivariable analysis, 4 showed substantial and statistically significant associations with pressure ulcer healing. These included ulcer stage, bed mobility, absence of paraplegia or quadriplegia, and absence of end-stage illness. Young age (ie, age < 65), male sex, urinary continence, intact bed transfer ability, and history of a resolved ulcer had a greater than 10% increase in odds of healing. We felt that these were important predictors of healing and therefore included them in the model as discussed in the Methods section. Cognitive status, diabetes, peripheral vascular disease, and congestive heart failure did not have a substantial association with the outcome. Among the other predictors that did not meet statistical significance, we excluded recent hip fracture, change in weight, presence of edema, and fever because we did not feel that they were important predictors of pressure ulcer healing beyond the predictors already included. Cancer and low BMI (ie, BMI < 19) counterintuitively predicted increased healing, although the results were

Table 3

Probability of Healing Divided Into Deciles for Derivation and Validation Samples Showing Expected Rates of Pressure Ulcer Healing from the Logistic Regression Model and the Observed Rates

	Derivation Set		Validation Set	
Predicted Decile	Expected Rate (%)	Observed Rate (%)	Expected Rate (%)	Observed Rate (%)
1	27.3	28.9	26.2	32.5
2	38.4	41.2	38.4	37.9
3	46.0	41.2	45.6	37.4
4	53.7	50.4	53.1	44.8
5	58.3	60.2	58.3	56.1
6	63.9	66.8	63.9	66.5
7	66.4	64.1	67.0	69.5
8	68.7	69.0	69.9	69.4
9	71.8	70.3	72.9	70.5
10	75.2	76.6	76.0	67.0

Table 4

Results from Multivariable Logistic Regression Model Fit to All Data With Outcome of Healing of Pressure Ulcers

Variable Names	OR Logistic	OR Logistic 95% CI
Age <65 y	1.08	0.83-1.39
Male sex	1.15	0.99-1.33
Stage 2 ulcer (compared to stage 4)	3.18	2.62-3.87
Stage 3 ulcer (compared to stage 4)	1.23	0.98-1.54
Intact bed mobility	1.56	1.22-2.00
Intact transferring from bed	1.21	0.94-1.56
Urinary continence	1.09	0.89-1.32
Bowel continence	1.12	0.93-1.36
Absence of paraplegia and quadriplegia	3.17	1.83-5.48
Absence of end-stage illness	1.45	1.06-2.00
History of resolved ulcer	1.13	0.95-1.34

Abbreviations: OR, odds ratio; CI, confidence interval.

not statistically significant. We excluded these variables as discussed.

In the sensitivity analysis looking at race, white race (compared with nonwhite race) was not significantly associated with healing (OR = 0.98; 95% CI = 0.81-1.19). Therefore, we did not include race in subsequent analyses. Results from sensitivity analysis, including duration of follow-up time, revealed that a longer than average duration (ie, 110-135 days) did not have a significant association with pressure ulcer healing; hence, duration was not included in the model.

In the derivation set, the C statistic for the model was 0.68, and the Hosmer-Lemeshow χ^2 was 6.50, (P = .59), indicating acceptable calibration. In Table 3, we report the expected and observed rates by deciles of model-predicted probability of healing. There are only isolated deciles in which the expected rate did not match the observed rate. The model was able to identify individuals with high and low model-predicted probability of ulcer healing, with close to a 3-fold difference in observed rate between the highest and lowest decile. A model with ulcer stage alone had a C statistic of 0.63.

In the validation set, the C statistic for the model was 0.66, and the Hosmer-Lemeshow χ^2 was 22.1 (P = .004). The model did not perform as well (>5% difference in observed and expected) in deciles 1, 3, 4, and 10 of predicted probability in this validation set (Table 3).

A final list of the associations between the variables and the outcome fit to the entire set of ulcer episodes appears in Table 4. A stage 2 ulcer predicted more healing compared with a stage 4 ulcer, (OR = 3.18; 95% CI = 2.62-3.87). Having some or full mobility in bed predicted more healing than having no mobility in bed (OR = 1.56; 95% CI = 1.22-2.00). Absence of paraplegia or quadriplegia predicted more healing (OR = 3.17; 95% CI = 1.83-5.48). Absence of end-stage illness also predicted more healing (OR = 1.45; 95% CI = 1.06-2.00). The C statistic for the entire dataset was 0.67, and Hosmer-Lemeshow χ^2 was 3.52, (P = .89).

Facility Performance

Nursing homes differed considerably in case mix as described by their expected rates of pressure ulcer healing. Among the 62 (of 110) nursing homes with 30 or more ulcer episodes, the expected range for healing was 47% to 65%. The differences were significant (P < .001) on test with ANOVA. These nursing homes also differed in their observed rates of healing, ranging from 23% to 74% (P < .001).

Judgments of nursing home performance varied depending on whether we considered unadjusted or risk-adjusted performance measures. Figure 1 depicts the performance profile for the 10 top-performing and 10 bottom-performing facilities ordered by decreasing observed rate. In the top 10 facilities, only G and H performed better than their acceptability

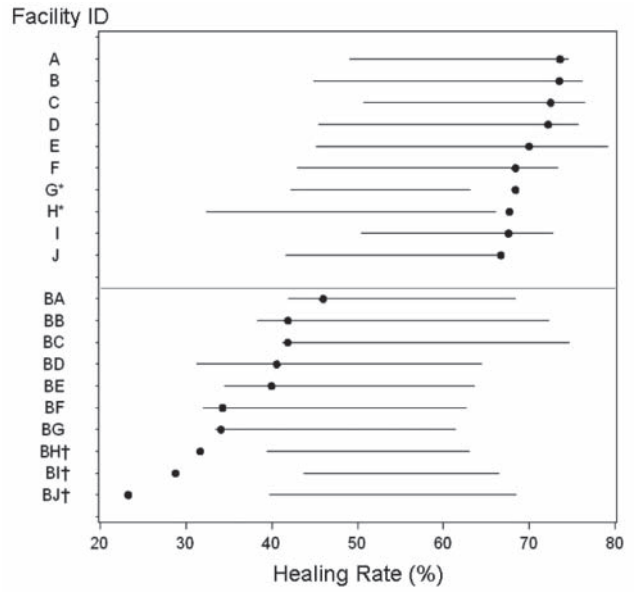


Figure 1. Performance profiles for facilities in top and bottom of performance.

• Observed Rate ——— Acceptability Interval (95% CI of Expected Rate)

* In these facilities, the observed rate of performance fell above the acceptability interval indicating top outlier status. Note that facilities with higher observed rates (A through F) would have been considered better performers had we not determined outlier status.
† In these facilities, the observed rate fell below the acceptability interval indicating bottom outlier status. These facilities were also the lowest performers by observed rate.

intervals. We would have considered facilities A through F as better performers than G and H if we had not considered outlier status. Judgments about the bottom performers did not differ whether we considered observed rates or the outlier status. Overall, there were 2 top outliers and 3 bottom outliers.

DISCUSSION

Using data from the MDS, we developed a risk-adjustment model with credible predictors of pressure ulcer healing in nursing home residents. The model had moderately good discriminatory capacity, identifying individuals with a 3-fold difference in observed rate between the lowest and highest deciles of model-predicted risk. This discriminatory ability was maintained in the validation phase of our analysis. Use of risk-adjusted outcomes based on this model was important in making judgments about nursing home performance.

Establishing accurate indicators of nursing home performance has become an important challenge in health services research. Risk-adjusted outcomes can serve as such indicators, yet controversy exists as to the extent of their use. Some investigators have argued against the use of risk adjustment in the long-term care setting because certain predictors, such as functional status, will vary over time and may be influenced by prior poor care.[13] Although inclusion of such predictors requires special attention, one must still account for the fact that different providers see very different populations of patients. Risk adjustment has been shown to alter judgments of provider performance in the current and previous work in the long-term care setting.[14-17] The controversy heralds the need to examine multiple indicators to comprehensively assess the quality of individual facilities.

There are only a handful of investigations that have attempted to delineate the predictors of pressure ulcer healing. Comparisons with previous models are limited given the variation in the outcome assessed and/or process of care predictors included. Using percent reduction in ulcer size, Kramer et al studied 109 patients with pressure ulcers and found that patient and ulcer characteristics explained 19% of the variation in healing.[5] Mukamel studied deterioration in ulcer level as a quality indicator.[1] Her model obtained good discrimination, a C statistic of 0.75, with acceptable calibration. This last model looked at both skilled nursing facility and health-related facility residents with a separate predictor included if the subject was in one type of facility or the other. The model also included the use of physical restraints. Similarly, Berlowitz et al published a study of pressure ulcer healing in 819 veterans using the VA Patient Assessment File (PAF) and included a predictor related to receipt of rehabilitation services.[2] Also, the C statistic in that model was higher (0.73) than that in the current study. The National Pressure Ulcer Long-Term Care Study found that resident characteristics explained little of the variation in the change in ulcer area.[18] Their model contained few patient variables and many treatment variables, which limits comparison with the current study. We chose a limited focus, including only the predictors that represent inherent risk of healing and not process of care, as is appropriate for a risk-adjustment model.

The variation in facility performance as expressed by unadjusted and risk-adjusted measures suggests that risk adjustment mattered in the current study. Specifically, the 2 top outliers would not have

been recognized as the top performers without risk adjustment. Confirming earlier work, simply looking at the observed rates of healing would distort the performance profile of the evaluated nursing homes.[3]

Although the model developed in this study is most useful in making judgments about quality of individual nursing facilities, the predictors that comprised the model offer some prognostic information for clinicians. Stage of ulcer explained the largest amount of variability in the rate of healing. Mobility and absence of end-stage illness predicted healing consistent with the above-mentioned VA PAF study. In contrast, urinary continence did not have a significant association with healing in the current study. The previous studies which identified urinary continence as a predictor may not have controlled adequately for variables, such as mobility, which are better characterized in the more clinically oriented MDS. Predictors of pressure ulcer formation that were reported by Berlowitz et al using MDS, such as diabetes, did not show the same level of association with pressure ulcer healing.[4] Paraplegia and quadriplegia were highly associated with healing independent of mobility in bed. This probably reflects a lack of sensory perception and/or an additional level of immobility that even the MDS mobility predictors do not capture.

Limitations include the loss to follow-up from discharge, death, or transfer. Gaming by facilities, wherein a facility transfers a patient whom it feels would not heal, may result in inflated healing rates. If a decision of a facility to transfer relied on predictors included in the current model, the effect of these predictors may have been mitigated and the overall model performance dampened. The lack of ulcer level information also was a limitation. We did not attempt to correlate the difference in performance in pressure ulcer healing with differences in processes of care used in the individual facilities. Although MDS contains information about certain process variables, such as use of a pressure ulcer relieving device or turning program, documentation of usage has not been found to correlate with actual usage or quality of care (such as when site visits were conducted).[19] Understanding how top outlier nursing homes obtain better outcomes and finding the accurate methods to measure these higher quality practices should be priorities for future research.

We did not consider other related outcomes, such as deterioration in ulcer status, when we profiled facilities. Ulcers that remain in the same stage could be scored more favorably than those that worsened. Because the number of ulcer episodes was relatively small at the facility level, we did not differentiate ulcer healing for ulcers that developed outside a facility from those that formed within it. Thereby, a facility could perform well in healing ulcers yet perform poorly in preventing them. The issues mentioned above reiterate the need to profile facilities in multiple dimensions. To provide a composite picture of pressure ulcer care, some investigators have profiled facilities by change in ulcer prevalence using a cohort design or point prevalence using cross-sectional analysis.[19-22] As an outcome, prevalence confounds the quality issue because admission of patients from outside facilities may change prevalence without informing consumers about the performance of the facility with respect to pressure ulcer care. Using prevalence also makes the process of feedback and quality improvement problematic because we cannot inform facilities about which processes they need to review, those related to healing or those related to prevention. We believe healing to be a cleaner outcome than prevalence but reiterate the need to consider multiple outcomes when evaluating the quality of a facility.

There is a movement toward adopting pay-for-performance strategies as a method to improve quality of care. This movement emphasizes the need for good data on quality of care provided. Our risk-adjusted model of pressure ulcer healing represents an attempt to develop one measure for evaluating nursing home performance. However, there is a need to validate the model externally and to develop a composite measure that addresses a broad range of outcomes.

Although our model focuses on the care of nursing home residents, a similar model can be applied to home care or other long-term care settings. The Online Analytical Statistical Information System (OASIS) was created to support the development of risk-adjustment models using data collected during home health care visits. Results from demonstration projects have shown that quality improvement programs incorporating risk-adjusted OASIS data can significantly improve patient outcomes, such as reducing hospitalization rates.[23] Future work should examine the performance of our model in alternate settings.

In conclusion, the risk-adjustment model we developed contains credible predictors of pressure ulcer healing. Use of risk-adjusted measures identified top-performing facilities that differed from those identified by unadjusted analysis. Future models should incorporate the predictors identified in this study (ie, ulcer stage, bed mobility, paraplegia

and quadriplegia, and end-stage illness). Pressure ulcer healing may be one of many indicators used to evaluate nursing home quality.

REFERENCES

1. Mukamel DB. Risk-adjusted outcome measures and quality of care in nursing homes. *Med Care.* 1997;35:367-385.
2. Berlowitz DR, Brandeis GH, Anderson J, Brand HK. Predictors of pressure ulcer healing among long-term care residents. *J Am Geriatr Soc.* 1997;45:30-34.
3. Berlowitz DR, Brandeis GH, Anderson JJ, et al. Evaluation of a risk-adjustment model for pressure ulcer development using the Minimum Data Set. *J Am Geriatr Soc.* 2001;49:872-876.
4. Berlowitz DR, Brandeis GH, Morris JN, et al. Deriving a risk-adjustment model for pressure ulcer development using the Minimum Data Set. *J Am Geriatr Soc.* 2001;49:866-871.
5. Kramer JD, Kearney M. Patient, wound, and treatment characteristics associated with healing in pressure ulcers. *Adv Skin Wound Care.* 2000;13:17-24.
6. Berlowitz DR, Ash AS, Brandeis GH, Brand HK, Halpern JL, Moskowitz MA. Rating long-term care facilities on pressure ulcer development: importance of case-mix adjustment. *Ann Intern Med.* 1996;124:557-563.
7. Allman RM, Goode PS, Patrick MM, Burst N, Bartolucci AA. Pressure ulcer risk factors among hospitalized patients with activity limitation. *JAMA.* 1995;273:865-870.
8. Perneger TV, Gaspoz JM, Rae AC, Borst F, Heliot C. Contribution of individual items to the performance of the Norton pressure ulcer prediction scale. *J Am Geriatr Soc.* 1282;46:1282-1286.
9. Davis PA, Corless DJ, Gazzard BG, Wastell C. Increased risk of wound complications and poor healing following laparotomy in HIV-seropositive and AIDS patients. *Dig Surg.* 1999;16:60-67.
10. Morandi E, Merlini D, Salvaggio A, Foschi D, Trabucchi E. Prospective study of healing time after hemorrhoidectomy: influence of HIV infection, acquired immunodeficiency syndrome, and anal wound infection. *Dis Colon Rectum.* 1140; 42:1140-1144.
11. Robinson PG, Cooper H, Hatt J. Healing after dental extractions in men with HIV infection. *Oral Surg, Oral Med, Oral Pathol.* 1992;74:426-430.
12. Sun GW, Shook TL, Kay GL. Inappropriate use of bivariable analysis to screen risk factors for use in multivariable analysis. *J Clin Epidemiol.* 1996;49:907-916.
13. Zimmerman DR, Karon SL, Arling G, et al. Development and testing of nursing home quality indicators. *Health Care Financ Rev.* 1995;16:107-127.
14. Arling G, Karon SL, Sainfort F, Zimmerman DR, Ross R. Risk adjustment of nursing home quality indicators. *Gerontologist.* 1997;37:757-766.
15. Mukamel DB, Brower CA. The influence of risk adjustment methods on conclusions about quality of care in nursing homes based on outcome measures. *Gerontologist.* 1998;38:695-703.
16. Porell F, Caro FG. Facility-level outcome performance measures for nursing homes. *Gerontologist.* 1998;38:665-683.
17. Rosen AK, Berlowitz DR, Anderson JJ, Ash AS, Kazis LE, Moskowitz MA. Functional status outcomes for assessment of quality in long-term care. *Int J Qual Health Care.* 1999; 11:37-46.
18. Bergstrom N, Horn SD, Smout RJ, et al. The National Pressure Ulcer Long-Term Care Study: outcomes of pressure ulcer treatments in long-term care. *J Am Geriatr Soc.* 2005;53:1721-1729.
19. Bates-Jensen BM, Cadogan M, Osterweil D, et al. The Minimum Data Set pressure ulcer indicator: does it reflect differences in care processes related to pressure ulcer prevention and treatment in nursing homes? *J Am Geriatr Soc.* 2003;51:1203-1212.
20. Newland PK, Wipke-Tevis DD, Williams DA, Rantz MJ, Petroski GF. Impact of pain on outcomes in long-term care residents with and without multiple sclerosis. *J Am Geriatr Soc.* 2005;53:1490-1496.
21. Bours GJ, Halfens RJ, Berger MP, Huijer Abu-Saad H, Grol RT. Development of a model for case-mix adjustment of pressure ulcer prevalence rates. *Med Care.* 2003;41:45-55.
22. Casimiro C, Garcia-de-Lorenzo A, Usan L. Prevalence of decubitus ulcer and associated risk factors in an institutionalized Spanish elderly population. *Nutrition.* 2002; 18:408-414.
23. Shaughnessy PW, Hittle DF, Crisler KS, et al. Improving patient outcomes of home health care: findings from two demonstration trials of outcome-based quality improvement. *J Am Geriatr Soc.* 2002;50:1354-1364.

The Impact of Trauma Exposure and Post-Traumatic Stress Disorder on Healthcare Utilization Among Primary Care Patients

Anand Kartha, MD, MSc,† Victoria Brower, MPH,‡ Richard Saitz, MD, MPH,*§¶ Jeffrey H. Samet, MD, MA, MPH,*‖ Terence M. Keane, PhD,†#** and Jane Liebschutz, MD, MPH*‖*

Background: Trauma exposure and post-traumatic stress disorder (PTSD) increase healthcare utilization in veterans, but their impact on utilization in other populations is uncertain.

Objectives: To examine the association of trauma exposure and PTSD with healthcare utilization, in civilian primary care patients.

Research Design: Cross-sectional study.

Subjects: English speaking patients at an academic, urban primary care clinic.

Measures: Trauma exposure and current PTSD diagnoses were obtained from the Composite International Diagnostic Interview. Outcomes were nonmental health outpatient and emergency department visits, hospitalizations, and mental health outpatient visits in the prior year from an electronic medical record. Analyses included bivariate unadjusted and multivariable Poisson regressions adjusted for age, gender, income, substance dependence, depression, and comorbidities.

Results: Among 592 subjects, 80% had ≥1 trauma exposure and 22% had current PTSD. In adjusted regressions, subjects with trauma exposure had more mental health visits [incidence rate ratio (IRR), 3.9; 95% confidence interval (CI), 1.1–14.1] but no other increased utilization. After adjusting for PTSD, this effect of trauma exposure was attenuated (IRR, 3.2; 95% CI, 0.9–11.7). Subjects with PTSD had more hospitalizations (IRR, 2.2; 95% CI, 1.4–3.7), more hospital nights (IRR, 2.6; 95% CI, 1.4–5.0), and more mental health visits (IRR, 2.2; 95% CI, 1.1–4.1) but no increase in outpatient and emergency department visits.

Conclusions: PTSD is associated with more hospitalizations, longer hospitalizations, and greater mental healthcare utilization in urban primary care patients. Although trauma exposure is independently associated with greater mental healthcare utilization, PTSD mediates a portion of this association.

Key Words: PTSD, trauma, primary care, hospitalization, utilization

(*Med Care* 2008;46: 388–393)

Trauma exposure can be described as an unexpected, violent encounter with nature, technology, or humankind.[1] Post-traumatic stress disorder (PTSD) is a syndrome after a significant trauma exposure characterized by persistent, recurrent symptoms including intrusive recollections, avoidance of disturbing stimuli and increased arousal, which cause significant functional impairment.[2]

Healthcare utilization data are essential for estimating the impact of specific medical conditions and for informing healthcare planning and policy.[3] Prior studies suggest that trauma exposure and PTSD have considerable impact on healthcare utilization and costs.[1,3–5] Most of this research has focused on male veterans and female sexual assault victims and reports increased healthcare utilization for mental and physical problems.[1,5–7] These findings may not be generalizable to other populations as the services available to veterans such as specialized PTSD clinics and disability compensation may differ from those available to civilians with PTSD.[3,7]

Trauma exposure and PTSD are common in primary care settings for nonveterans and veterans alike.[1,5,8–10] However, few studies have examined healthcare utilization patterns in this population.[2,11–16] The studies among veterans have generally shown increased use of mental healthcare services but conflicting findings in the use of services for physical health problems. As noted by Elhai and others in a recent review of healthcare utilization studies in trauma survivors, these studies have been limited by methodological concerns including not using diagnostic measures for PTSD,[2,13–16] small sample sizes,[2,11–14,16] lack of adjustment

From the *Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Boston Medical Center and Boston University School of Medicine; †Veterans Affairs Boston Healthcare System; ‡Massachusetts Behavioral Health Partnership, Boston, Massachusetts; §Department of Epidemiology, ¶Youth Alcohol Prevention Center, Boston University School of Public Health; ‖Department of Social and Behavioral Sciences, #Department of Psychiatry, Boston University School of Medicine; and **National Center for Post Traumatic Stress Disorder; all in Boston, Massachusetts.

Supported in part by a Generalist Physician Faculty Scholar Award from the Robert Wood Johnson Foundation, Princeton, NJ (RWJF #045452), and by the National Institute on Drug Abuse, Bethesda, MD (K23 DA01665), both awarded to Dr. Liebschutz.

Portions of this work were presented at the annual meeting of the Society of General Internal Medicine, May 2005, New Orleans, LA; at the annual meeting of the College on Problems of Drug Dependence, June 2005, Orlando, FL, and at the annual meeting of the American Public Health Association, November 2004, Washington, DC.

Reprints: Anand Kartha, MD, MSc, VA Boston Healthcare System, 1400 VFW Parkway, Mail Stop 111, West Roxbury, MA 02132. E-mail: anand.kartha@med.va.gov.

ISSN: 0025-7079/08/4604-0388

for potential confounders of utilization,[2,11,13,14,16] and use of self-reported utilization measures,[11,12,14] which can be less accurate.[3,17] Data collected from electronic medical records (EMR) are considered to be more valid than self-reported utilization data.[3] Because these studies have mostly been conducted in populations with higher socioeconomic status, they may not be applicable to residents of urban, economically disadvantaged areas who may be both at greater risk for trauma exposure as well as more vulnerable to the effects of trauma.[12,15] Finally, most studies to date have not clarified whether health service use is independently affected by trauma exposure, subsequent development of PTSD, or both.

This study's objective was to examine the independent associations of both trauma exposure and PTSD with healthcare utilization in a civilian, urban primary care population and to examine any mediating role of PTSD in the relationship between trauma exposure and utilization. We hypothesized that urban primary care patients with trauma exposure or PTSD would have more nonmental health outpatient and emergency department visits, hospitalizations and mental health outpatient visits. We further hypothesized that any increases in utilization associated with trauma exposure would be explained in part by the presence of PTSD.

METHODS

The current analyses were conducted as part of a cross-sectional study of PTSD and its comorbidities in primary care. We interviewed a sample of primary care patients at the outpatient department of an urban university-affiliated, safety-net hospital to examine overall prevalence of traumatic exposure and select behavioral health outcomes in addition to PTSD, including: major depression, substance dependence and chronic pain. A detailed description of the main study methods can be found elsewhere[9]; we summarize methods relevant to the current analyses below.

Subjects

Patients presenting to the primary care (Internal Medicine and Family Medicine) clinics at an urban, academic safety-net medical center were eligible if they spoke English, were between the ages of 18 and 65 years and had a scheduled appointment with a primary care clinician. Patients were excluded if they could not be interviewed alone or if they did not seem to understand the study. After the first 509 consecutive patients were enrolled meeting all inclusion criteria, we limited enrollment to a "subsample" of eligible patients oversampled for alcohol and drug use and irritable bowel symptoms to permit preplanned subgroup analyses ("subsample" = 98 patients). The current analyses used the entire sample of 607. Boston University Medical Center's Institutional Review and HIPAA Privacy Review Boards approved the study. A Certificate of Confidentiality was obtained from the National Institutes of Health.

Assessments

Interviewers screened consecutive patients arriving at the primary care clinic and eligible patients were asked to participate in an interview about stress and health after obtaining informed consent. The interview included demographic questions, the Composite International Diagnostic Interview (CIDI) version 2.1 PTSD module,[18] the Chronic Pain Definitional Questionnaire,[19] the Patient Health Questionnaire (PHQ) modules measuring depression diagnoses (major and other depression in past 2 weeks)[20] and the CIDI-Short Form (CIDI-SF) modules for drug and alcohol dependence in the past 6 months.[18] Enrollment took place from February 2003 to September 2005.

Using standardized data forms, trained researchers reviewed patient EMR to collect nonmental health outpatient and emergency department (ED) visits, hospitalizations and mental health outpatient visits. The EMR registration history records every single visit in all inpatient and outpatient locations of our comprehensive medical system, which provides most of our population's healthcare.

Study variable selection was guided by Andersen's behavioral model of health services use, a well validated and extensively used conceptual framework for healthcare utilization.[21] In this model, subjects use healthcare resources depending on their medical needs, which in turn may be influenced by predisposing and enabling factors. Need refers to actual illness, which is the most important determinant of healthcare use and we conceptualized trauma exposure and PTSD as need factors in our models. Comorbid medical illness and depression were also characterized as need factors because they could also influence healthcare use in patients. Predisposing factors influence the propensity of an individual to seek medical attention and we characterized age, gender, and substance dependence as predisposing factors. Finally, enabling factors enhance or impede use of healthcare; we characterized income, for example, as an enabling factor. Although health insurance status is often considered as an enabling factor, we did not include insurance status in our model as >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").

Independent variables were: (1) trauma exposure defined using 11 questions (9 questions on specific types of traumatic events and 2 open ended questions on exposure) from the CIDI, a well validated and reliable diagnostic interview developed by the World Health Organization for diagnosing mental disorders based on DSM-IV criteria.[18] We first characterized trauma exposure as a dichotomous variable (none vs. one or more). We also characterized number of types of traumatic events as a class variable (0, 1–2, 3–4, ≥5 types of traumas) to assess a "dose-response" relationship between trauma exposure and utilization. The classes approximately divided the sample population into 4 quartiles; (2) diagnosis of current PTSD (past 12 months), characterized as a dichotomous variable, obtained by the CIDI.[18] *Dependent (outcome) variables* were counts of nonmental health outpatient visits involving any direct clinician contact, ED visits, inpatient hospitalizations for any reason, and mental health outpatient visits, for the 12 months before the research interview. We also counted the inpatient nights spent during hospitalizations as a proxy measure for the amount of inpatient resources used. Multiple visits on the same day to different clinicians were counted as discrete outcomes. Visits

for diagnostic testing were excluded. *Covariates* used in adjusted models included age, gender, annual income (≤ or >$20,000), substance dependence (drug and/or alcohol dependence in the past 6 months by CIDI-SF),[18] depression (major and/or other depression by PHQ),[20] and medical comorbidity measured by the Charlson Index (obtained from the diagnostic information in the EMR. Possible range of scores was 0–33. Higher scores indicate greater medical comorbidity).[22]

Statistical Methods

All analyses used 2-tailed tests with a P value of ≤0.05 considered to indicate statistical significance. Descriptive and bivariate unadjusted analyses included χ^2 tests to evaluate differences in proportions, t tests for differences between means for parametric data and Wilcoxon rank sum tests for differences between medians for skewed data. Multivariable adjusted analyses, to determine the independent associations between trauma exposure and PTSD with utilization included Poisson regression models accounting for overdispersion to obtain incidence rate ratios (IRR).[23] We used a Poisson regression because the utilization dependent variables are count measures with a skewed, nonparametric distribution, and therefore standard parametric approaches like linear regression are not statistically appropriate.[23] To determine whether PTSD mediated the relationship between trauma exposure and utilization, we fit 2 adjusted Poisson regression models, the first including trauma exposure and utilization outcome variables and the second adding the PTSD variable. To test whether the subsample of participants recruited for substance use and irritable bowel syndrome influenced findings, we fit another model excluding the subsample. The results were essentially unchanged and the data are not presented here. All analyses were done using SAS software, version 8.0.

RESULTS

Of the 751 patients who met eligibility criteria for this study, 607 (81%) agreed to participate. The 144 subjects (19%) who refused did not differ significantly from study subjects in terms of age, gender, or race. Nonparticipants were more likely to be born outside the United States (41% vs. 21%, P = 0.001). The 98 subjects in the "subsample" did not differ significantly from other subjects in age, race, marital status, employment, income, and prevalence of PTSD. They had a greater prevalence of major depression (37% vs. 21%, P = 0.001) and trauma exposure (88% vs. 79%, P = 0.04). Medical record information was not available for 15 of the 607 subjects, leaving 592 subjects for analysis. See Table 1 for participant characteristics.

Trauma Exposure

Eighty percent of subjects had one or more trauma exposures. Compared with subjects with no trauma exposure, subjects with trauma exposure were significantly more likely to be male (51% vs. 38%), unmarried (87% vs. 75%), have substance dependence (21% vs. 5%), and depression (48% vs. 33%).

TABLE 1. Characteristics of Primary Care Patients

	Total, % (N = 592)	PTSD, % (N = 133)	No PTSD, % (N = 459)	P
Age (yr), mean ± SD	41.6 ± 11.6	40.6 ± 10.7	41.9 ± 11.8	0.3
Female gender	51	62	48	0.007
Ethnicity*				0.8
Black	59	56	60	
White	19	18	19	
Hispanic	8	10	8	
Other	14	16	14	
Married	15	11	16	0.10
Education*				0.2
<High school	25	30	23	
High school	34	34	34	
College	41	36	43	
Annual Income*				<0.0001
≤$20,000	50	68	46	
>$20,000	50	33	55	
Drug and/or alcohol dependence past 6 mo*	18	24	16	0.04
Major and/or other depression	45	71	37	<0.0001
Comorbidity—mean ± SD*†	0.66 ± 1.29	0.66 ± 1.36	0.66 ± 1.26	0.9

*No. patients for whom data not available: ethnicity = 1, education = 2, income = 18, substance dependence = 4, comorbidity = 1.
†As measured by the Charlson Comorbidity Index. Range of scores in this group of subjects was 0–9.

TABLE 2. Unadjusted Association of PTSD With 12-Month Utilization (N = 592)

	PTSD (N = 133)	No PTSD (N = 459)	P
Outpatient visits*	9.16 ± 9.74 (7)	8.35 ± 9.95 (5)	0.10
ED visits*	1.88 ± 3.28 (1)	1.41 ± 2.76 (1)	0.2
Hospitalizations*	0.43 ± 1.32 (0)	0.18 ± 0.57 (0)	0.05
Inpatient nights*	2.99 ± 14.88 (0)	1.01 ± 4.34 (0)	0.08
Mental health visits*	1.41 ± 4.74 (0)	0.5 ± 2.39 (0)	<0.0001

*Mean ± SD (median).

In bivariate unadjusted analyses, subjects with trauma exposure had more mental health visits (mean 0.83 vs. 0.18, median 0 vs. 0, P = 0.005) and emergency department visits (mean 1.55 vs. 1.37, median 1 vs. 0, P = 0.03) compared with subjects with no trauma exposure. In multivariable adjusted analyses (Table 3),), subjects with trauma exposure had 3.90 times more mental health visits in 12 months compared with patients with no trauma exposure (IRR, 3.90; 95% CI, 1.08–14.14). This association was attenuated and no longer statistically significant after PTSD was added as a covariate in this model (IRR, 3.16; 95% CI, 0.85–11.69). Subjects with trauma exposure did not have more outpatient visits (IRR, 1.10; 95% CI, 0.89–1.40), ED visits (IRR, 1.07; 95% CI, 0.71–1.60) or hospitalizations (IRR, 0.94; 95% CI, 0.49–1.77).

TABLE 3. Incidence Rate Ratios of PTSD and Trauma Exposure With Prior 12-Month Utilization*

	PTSD† (95% CI)	Trauma Exposure‡ (95% CI)
Outpatient visits	1.04 (0.84–1.28)	1.10 (0.89–1.40)
ED visits	1.07 (0.75–1.54)	1.07 (0.71–1.60)
Hospitalizations	2.22 (1.35–3.67)	0.94 (0.49–1.77)
Inpatient nights	2.62 (1.38–4.99)	0.84 (0.36–1.95)
Mental health visits	2.15 (1.14–4.06)	3.90 (1.08–14.14)

*Poisson regressions adjusted for age, gender, income, substance dependence, depression, and comorbidity.
†Total no. subjects available for analyses = 570; reference group for IRR is no PTSD.
‡Total no. subjects available for analyses = 577; reference group for IRR is no trauma exposure.

We found a similar pattern of results when we studied trauma exposure severity by characterizing it as a 4-class variable (0, 1–2, 3–4, ≥5 types of trauma). For example, subjects with ≥5 types of trauma had significantly more mental health visits (IRR, 3.19; 95% CI, 1.22–8.34) when compared with subjects with no traumas, but this association was attenuated and no longer significant after PTSD was added as a covariate in this model (IRR, 2.06; 95% CI, 0.69–6.20).

PTSD

Among the 592 subjects, 133 (22%) had current PTSD. Mean duration of PTSD was 11.9 years (median = 7.9 years) and 94% of subjects had PTSD symptoms for 11 months or more. Compared with subjects without PTSD, subjects with PTSD were significantly more likely to be female (62% vs. 48%), to have an annual income less than or equal to $20,000 (68% vs. 46%), to meet criteria for substance dependence (24% vs. 16%) and depression (71% vs. 37%) (Table 1). In bivariable unadjusted analyses (Table 2), subjects with PTSD had significantly more hospitalizations (mean number 0.43 vs. 0.18, median 0 vs. 0, $P = 0.05$) and mental health visits (mean 1.41 vs. 0.50, median 0 vs. 0, $P < 0.0001$) in the prior 12 months than did subjects without PTSD. Although not reaching conventional levels of statistical significance, subjects with PTSD tended to spend more nights in the hospital (mean 2.99 vs. 1.01, median 0 vs. 0, $P = 0.08$) and higher outpatient (mean 9.16 vs. 8.35, median 7 vs. 5, $P = 0.10$) and emergency department visits (mean 1.88 vs. 1.41, median 1 vs. 1, $P = 0.16$). In multivariable adjusted analyses (Table 3), compared with subjects without PTSD, subjects with PTSD had 2.22 times more hospitalizations (IRR = 2.22, 95% CI = 1.35–3.67) and spent a greater number of nights in the hospital (IRR = 2.62, 95% CI = 1.38–4.99). They had 2.15 times more mental health visits (IRR = 2.15, 95% CI = 1.14–4.06). They did not have more outpatient visits (IRR = 1.04, 95% CI = 0.84–1.28) or ED visits (IRR = 1.07, 95% CI = 0.75–1.54).

DISCUSSION

Among urban primary care patients, PTSD is associated with greater healthcare use: both mental health visits and hospitalizations. Unexpectedly, trauma exposure by itself was not associated with increased utilization, apart from mental health visits, a finding which was attenuated after adjusting for PTSD.

The burden of depression alone on increased utilization of healthcare resources is well characterized. However, there is some controversy as to the incremental burden of anxiety disorders such as PTSD in primary care patients.[24] Our study confirms the increased mental healthcare use by PTSD patients seen in most prior studies in civilian primary care populations.[2,11,12,14,15] Although this is not surprising, our study highlights the additional burden of mental illness due to PTSD that persists even after accounting for the various comorbid medical conditions and depression that affects this population.

Our finding of greater use of inpatient but not of ED or outpatient resources contributes to the few prior studies in this area. These studies have found inconsistent results when reporting about healthcare use for physical problems in civilian primary care patients.[2,11–16] Stein et al found greater self-reported hospitalizations by patients with PTSD in a bivariate unadjusted analysis of approximately 18 primary care patients with PTSD compared with 74 patients without psychiatric disease, whereas Walker et al did not find increased hospitalization costs in a group of female HMO patients with PTSD.[14,15] However, the higher socioeconomic status of their HMO population may reflect healthier patients with more resources to avoid hospitalization.[15] Although Stein et al found increased ED use, the larger study by Walker et al did not find an association between PTSD and ED costs, which is consistent with the nonsignificant association we found between PTSD and ED visits.[14,15] To our knowledge, no other additional studies besides these 2 have examined the association between PTSD and use of inpatient and ED resources in nonveteran primary care patients.[14,15] A few studies have looked at nonmental health outpatient utilization in PTSD, but all of these studies had methodological limitations.[2,11,14–16] We did not find the increased use of outpatient resources seen in these studies. This may also reflect the overall high use of outpatient care by our patients, particularly because a study entry criterion was use of primary care. Taken collectively, our data suggest a pattern of increased use of healthcare resources for both mental and physical illness in urban primary care patients with PTSD, similar to the increased utilization seen in veterans.

There may be a number of potential pathways that explain increased utilization of health services for physical illness by some patients with PTSD. PTSD and trauma survivors have a higher prevalence of cardiac, digestive, musculoskeletal, nervous system, endocrine, and other physical illness; abnormal hypothalamic pituitary and sympathetic-adrenal-medullary axes; higher circulating T-cell counts; and immunoglobulin-M levels with lower cortisol and dehydroepiandrosterone levels, all of which strongly suggest an underlying deleterious biologic response to stress.[1,25] One could hypothesize that individuals can have varying psychologic responses to trauma such as high risk behaviors, medically unexplained somatic symptoms and marked physical and emotional functional impairment, all of which could increase their risk of hospitaliza-

tion.[1,24,26] Clinicians in turn commonly fail to recognize PTSD or may mistake somatization symptoms as manifestations of physical illness resulting in unnecessary tests and hospitalizations.[7–9] Notably, these responses can be sustained—over one-third of patients with PTSD have persistent symptoms after 10 years.[5]

Importantly, trauma exposure alone was not associated with any utilization for physical problems, an unexpected finding. In the one area in which we did find increased utilization with trauma exposure, that of mental healthcare, PTSD attenuated this association. Given that 80% of our study subjects were exposed to at least one type of trauma, it is possible that any other differences in utilization secondary to trauma may be masked in this highly traumatized population. Prior studies have consistently shown greater utilization with specific traumas such as criminal or sexual assault; whereas we did not attempt to distinguish between different types of trauma, suggesting that all trauma exposures may not be equal.[5] Finally, other factors besides exposure to trauma alone may be necessary to influence healthcare seeking behavior. Although some trauma survivors may be resilient and avert the deleterious effects of trauma, others may be more vulnerable; we did not attempt to characterize these differences. Similarly, development of PTSD after trauma exposure may be one of the pathways mediating healthcare use.[16] This is supported by our findings that adjusting for the effect of PTSD attenuated the association of trauma exposure with increased mental healthcare utilization.

Limitations of this study include its cross-sectional design that precludes determination of causality of the association between PTSD or trauma exposure and utilization. However, it seems less likely that an increased use of medical resources would result in significant PTSD or trauma. Although, like almost all other similar studies we could not prospectively assess utilization in our patients, prior utilization is a powerful predictor of future utilization and may suggest continued greater use of resources by patients with PTSD.[15,27] The prevalence of PTSD was higher in this primary care population than noted in previous community studies. Although this may limit the generalizability of these results to other settings, it is certainly applicable to other urban hospitals that provide a large proportion of care for medically underserved populations with similar illness burdens. Although our EMR does not capture utilization outside the medical center, our comprehensive medical care system provides the majority of care for our population. Additionally, all our subjects had an established primary care provider at our medical center making it more likely they would get their healthcare within our system.

The relationship between trauma/PTSD and utilization has historically been studied in victims of combat and sexual assault, whereas almost no prior studies have been in urban, minority, disenfranchised community populations with a heavy burden of trauma like our subjects.[1,4] Although it may be expected that trauma and PTSD may not have an incremental effect on utilization in such populations given that participants have multiple other reasons for high utilization, our findings suggest that even on top of poor social conditions, PTSD impacts healthcare utilization. Furthermore, to our knowledge, no other study has addressed the serious methodological concerns of prior studies in this population—including lack of diagnostic criteria for PTSD, failure to adjust for medical and psychosocial factors known to affect utilization, and self-reported utilization data.[3] We confirmed the robustness of our trauma exposure findings by studying the effects of both the presence of trauma as well as the severity of trauma on utilization. Additionally, as pointed out by Hidalgo et al, other studies have not distinguished between PTSD and trauma exposure in the effect on healthcare utilization.[1] Our results suggest that future studies of utilization should account for development of PTSD in trauma survivors and that in poor urban clinical samples, trauma alone may not account for changes in utilization. This is also one of the largest studies to date to comprehensively study the patterns of healthcare utilization among primary care patients with PTSD in a nonveteran setting. Most prior community studies were done in HMO populations with higher socioeconomic status and our findings are applicable to other urban medical settings, which also have a large, similar trauma burden.[9,15] Our study thus contributes estimates of healthcare utilization secondary to trauma exposure in a population where trauma is extremely common, seldom recognized, and consequences are largely unknown.[1,4,9] Such estimates are critical to help policymakers and providers allocate and evaluate effectiveness of scarce healthcare resources.[1,4]

We conclude that in an urban, primary care population, PTSD is independently associated with a doubling of the number of hospitalizations and over twice the utilization of mental health resources. PTSD is not associated with more outpatient or ED visits. Although trauma exposure is associated with nearly four times the utilization of mental health resources, it seems that this effect may be partially mediated by PTSD. Trauma exposure alone does not seem to be associated with greater use of health resources for physical illness. Future studies of trauma survivors in health care settings should consider the use of available PTSD screening tools such as the PTSD Checklist to identify those who may benefit most from targeted allocation of resources.[28] Although the burden of depression on use of medical resources is well known, policymakers and clinicians are less aware of this significant impact of PTSD on use of medical services, particularly in nonveteran settings. Future, well-designed studies are required to prospectively determine both the mechanisms of how PTSD may contribute to utilization and if this additional utilization is appropriate. Effective treatment options now exist for PTSD[29]; thus, earlier efforts at detection and treatment in the primary care setting can be explored as a potential path to reduce both appropriate and inappropriate hospitalizations.

ACKNOWLEDGMENTS

We would like to acknowledge research assistance from Minga Claggett-Borne, Jessica Geier, Eric Holder, Lauren Kelly, Mary Reyes, Michael Rosa, Mary Benita Schickel, Pavan Sekhar and Jennifer Tran, statistical advice from Debbie Cheng, ScD and data management aid from Chris Lloyd-Travaglini.

REFERENCES

1. Hidalgo RB, Davidson JR. Posttraumatic stress disorder: epidemiology and health-related considerations. *J Clin Psychiatry*. 2000;61(Suppl 7):5–13.
2. Samson AY, Bensen S, Beck A, et al. Posttraumatic stress disorder in primary care. *J Fam Pract*. 1999;48:222–227.
3. Elhai JD, North TC, Frueh BC. Health service use predictors among trauma survivors: a critical review. *Psychol Serv*. 2005;2:3–19.
4. Kessler RC. Posttraumatic stress disorder: the burden to the individual and to society [review]. *J Clin Psychiatry*. 2000;61(Suppl 5):4–12; discussion 13–14.
5. Solomon SD, Davidson JR. Trauma: prevalence, impairment, service use, and cost [review]. *J Clin Psychiatry*. 1997;58(Suppl 9):5–11.
6. Calhoun PS, Bosworth HB, Grambow SC, et al. Medical service utilization by veterans seeking help for posttraumatic stress disorder. *Am J Psychiatry*. 2002;159:2081–2086.
7. Magruder KM, Frueh BC, Knapp RJ, et al. Prevalence of posttraumatic stress disorder in Veterans Affairs primary care clinics. *Gen Hosp Psychiatry*. 2005;27:169–179.
8. Lecrubier Y. Posttraumatic stress disorder in primary care: a hidden diagnosis [review]. *J Clin Psychiatry*. 2004;65(Suppl 1):49–54.
9. Liebschutz J, Saitz R, Brower V, et al. PTSD in urban primary care: high prevalence and low physician recognition. *J Gen Intern Med*. 2007;22:719–726.
10. Alim TN, Graves EM, Mellman TA, et al. Trauma exposure, posttraumatic stress disorder and depression in an African-American primary care population. *J Natl Med Assoc*. 2006;98:1630–1636.
11. Amaya-Jackson L, Davidson JR, Hughes DC, et al. Functional impairment and utilization of services associated with posttraumatic stress in the community. *J Trauma Stress*. 1999;12:709–724.
12. Switzer GE, Dew MA, Thompson K, et al. Posttraumatic stress disorder and service utilization among urban mental health center clients. *J Trauma Stress*. 1999;12:25–39.
13. Rosenberg HJ, Rosenberg SD, Wolford GL II, et al. The relationship between trauma, PTSD, and medical utilization in three high risk medical populations. *Int J Psychiatry Med*. 2000;30:247–259.
14. Stein MB, McQuaid JR, Pedrelli P, et al. Posttraumatic stress disorder in the primary care medical setting. *Gen Hosp Psychiatry*. 2000;22:261–269.
15. Walker EA, Katon W, Russo J, et al. Health care costs associated with posttraumatic stress disorder symptoms in women. *Arch Gen Psychiatry*. 2003;60:369–374.
16. Gillock KL, Zayfert C, Hegel MT, et al. Posttraumatic stress disorder in primary care: prevalence and relationships with physical symptoms and medical utilization. *Gen Hosp Psychiatry*. 2005;27:392–399.
17. Roberts RO, Bergstralh EJ, Schmidt L, et al. Comparison of self-reported and medical record health care utilization measures. *J Clin Epidemiol*. 1996;49:989–995.
18. World Health Organization. Composite International Diagnostic Interview (CIDI): Version 2.1. Geneva, Switzerland: World Health Organization; 1997.
19. Purves AM, Penny KI, Munro C, et al. Defining chronic pain for epidemiological research assessing a subjective definition. *Pain Clinic*. 1998;10:139–147.
20. 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*. 1999;282:1737–1744.
21. Andersen RM. Revisiting the behavioral model and access to medical care: does it matter? *J Health Soc Behav*. 1995;36:1–10.
22. 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;40:373–383.
23. Kianifard F, Gallo PP. Poisson regression analysis in clinical research. *J Biopharm Stat*. 1995;5:115–129.
24. Stein MB, Roy-Byrne PP, Craske MG, et al. Functional impact and health utility of anxiety disorders in primary care outpatients. *Med Care*. 2005;43:1164–1170.
25. Boscarino JA. Posttraumatic stress disorder and physical illness: results from clinical and epidemiologic studies. *Ann N Y Acad Sci*. 2004;1032:141–153.
26. Zatzick D. Posttraumatic stress, functional impairment, and service utilization after injury: a public health approach [review]. *Semin Clin Neuropsychiatry*. 2003;8(3):149–157.
27. Pakerson GR, Harrell FE Jr, Hammond WE, et al. Characteristics of adult primary care patients as predictors of future health service charges. *Med Care*. 2001;39:1170–1181.
28. Wilson JP, Keane TM. *Assessing Psychological Trauma and PTSD*. 2nd ed. New York, NY: Guilford Press; 2004.
29. Keane TM, Marshall AD, Taft CT. Posttraumatic stress disorder: etiology, epidemiology, and treatment outcome. *Ann Rev Clin Psychol*. 2006;2:161–197.

Willingness to Participate in Cancer Screenings: Blacks vs Whites vs Puerto Rican Hispanics

Ralph V. Katz, DMD, MPH, PhD, Cristina Claudio, PhD, Nancy R. Kressin, PhD, B. Lee Green, PhD, Min Qi Wang, PhD, and Stefanie Luise Russell, DDS, MPH, PhD

Background: *In the United States, blacks and Hispanics have lower cancer screening rates than whites have. Studies on the screening behaviors of minorities are increasing, but few focus on the factors that contribute to this discrepancy. This study presents the self-reported willingness by blacks, Puerto Rican Hispanics, and non-Hispanic whites to participate in cancer screenings in differing cancer screening situations.*

Methods: *The Cancer Screening Questionnaire (CSQ), a 60-item questionnaire, was administered via random-digit-dial telephone interviews to adults in three cities: Baltimore, Maryland; New York, New York; and, San Juan, Puerto Rico.*

Results: *The 1,148 participants in the CSQ study sample consisted of 355 blacks, 311 Puerto Rican Hispanics, and 482 non-Hispanic whites. Response rates ranged from 45% to 58% by city. Multivariable logistic regression analyses revealed that blacks and Puerto Ricans were often more likely (OR 2.0–3.0) and never less likely than whites to self-report willingness to participate in cancer screenings regardless of who conducted the cancer screening, what one was asked to do in the cancer screening, or what type of cancer was involved (with the exception of skin cancer where blacks, compared with whites, had an OR of 0.5).*

Conclusions: *The findings from this study provide evidence that blacks and Hispanics self-report that they are either as willing or more willing than whites to participate in cancer screening programs.*

From the Department of Epidemiology & Health Promotion at the New York University College of Dentistry, New York, New York (RVK, SLR); the Community Dentistry Section at the University of Puerto Rico School of Dentistry, San Juan, Puerto Rico (CC); the Department of General Internal Medicine at the Boston University Medical School and the Department of VA Health Services Research & Development Service at the US Department of Veterans Affairs Healthcare System, Boston, Massachusetts (NRK); the Office of Institutional Diversity at the H. Lee Moffitt Cancer Center & Research Institute, Tampa, Florida (BLG); and the Department of Public & Community Health at the University of Maryland School of Public Health, College Park, Maryland (MQW).

Submitted April 9, 2007; accepted November 27, 2007.

Address correspondence to Ralph V. Katz, DMD, MPH, PhD, Department of Epidemiology & Health Promotion, New York University College of Dentistry, 345 E. 24th Street, New York, NY 10012. E-mail: ralph.katz@nyu.edu

Abbreviations used in this paper: CSQ = Cancer Screening Questionnaire, TLP = Tuskegee Legacy Project.

Introduction

Cancer is the second-leading cause of death in the United States and affects people of all racial and ethnic groups. However, the rates of cancer occurrence and cancer death vary significantly between ethnic groups. For example, overall cancer incidence and mortality rates are higher for blacks compared with other groups. According to the National Cancer Institute in 2005,[1] the cancer incidence rate of blacks in the United States was 512.3 per 100,000, while for whites and Hispanics, it was 479.7 and 352.4 per 100,000, respectively. A similar pattern was observed regarding cancer death rates (African Americans = 248.1, whites = 195.3, and Hispanics = 135.2 per 100,000).

Differences by gender have also been observed. Black females, for example, have the highest incidence rates for colorectal cancer (56.0) and lung and

bronchus cancer (55.2) and the highest mortality rate for breast cancer (37.5). White females have the highest incidence for breast cancer (119.4) and the highest cancer death rate for lung and bronchus,[1] and Hispanic women have the highest incidence for cervical cancer (15.8 per 100,000).[1,2] Black men have the highest incidence and mortality rates for prostate, colorectal, and lung and bronchus cancers. Puerto Rican Hispanic males, living in either Puerto Rico or the United States, also have high incidence and mortality rates for prostate, colorectal, and lung and bronchus cancers.[1,3,4]

While the burden of cancer in minority populations is high, attempts to provide cancer screening examinations to minority populations have met with varied success at best.[5-9] The evidence is clear that early detection of cancer is one of the most effective means of lowering cancer mortality rates,[10,11] but it is also clear that blacks and Hispanics have lower cancer screening rates than whites in the United States.[5,6] Given that disadvantaged minority populations have less access to health care, cancer screening programs often provide the only means of early detection for a substantial proportion of minority populations.[12] The body of literature on reported past cancer screening behaviors of minorities has been expanding,[5-9,12-15] but fewer reports focus on the underlying factors of the willingness of minorities to participate in cancer screenings, especially compared with whites.[16] Most studies on factors related to minority participation in cancer screenings have been qualitative studies, usually utilizing focus group methodology that targets defining the concepts, language, and parameters of the issue.[17-28] Far fewer quantitative studies have investigated these factors in minority populations,[29,30] and only one study focused directly on the comparative question of whether minority populations are as willing as whites to participate in cancer screenings.[16]

A recent article focusing on the comparative willingness to participate in cancer screenings between Puerto Ricans living in San Juan, Puerto Rico, and Puerto Rican Americans living in New York, NY, reported that (1) those living in San Juan had a slightly but consistently higher self-reported willingness to participate in cancer screenings than those living in New York, and (2) the odds of San Juan Puerto Ricans participating in skin cancer screening, compared with New York Puerto Ricans were 3-fold higher in regard to participation in skin cancer screenings.[16]

The overall aim of the 3-City Cancer Screening Study was to address and understand a range of issues related to factors that determine whether blacks and Puerto Rican Hispanics, compared with non-Hispanics whites, would be willing to take part in cancer screening examinations. This understanding is critical to achieving early cancer detection goals in order to reduce the higher cancer mortality currently observed in minority populations. Specifically, this report focused on the self-reported willingness of blacks, Puerto Rican Hispanics, and non-Hispanic whites to participate in cancer screenings under differing cancer screening situations: (1) for differing site-specific types of cancer screening examinations, (2) for differing health care providers/institutions conducting the cancer screenings, and (3) for differing circumstances as to "what one is asked to do" as a part of that cancer screening examination.

Methods

The 3-City Cancer Screening Study was designed to administer the 60-item Cancer Screening Questionnaire (CSQ) via random-digit-dial telephone interviews to 900 adults aged 18 years and older in three cities: Baltimore, Maryland; New York, New York; and San Juan, Puerto Rico. The within-city recruitment goals were to recruit 300 blacks (150 in New York and 150 in Baltimore), 300 whites (150 in New York and 150 in Baltimore), and 300 Puerto Rican Hispanics (150 in New York and 150 in San Juan). The CSQ was developed in 2001 by a multidisciplinary, multi-university research team within the New York University Oral Cancer Research on Adolescent and Adult Health Promotion (RAAHP) Center, an Oral Health Disparities Center, funded from 2001-2009 by the National Institute of Dental and Craniofacial Research (NIDCR) at the National Institutes of Health (NIH). The CSQ addresses a range of issues related to the willingness of minorities to participate in cancer screening examinations, with a focus on whether minorities are more reluctant to participate in cancer screening examinations and, if so, the reasons for their reluctance.

An international opinion research corporation (Macro International, Inc, Burlington, Vermont) conducted the random-digit-dial survey using a computer-assisted telephone interviewing (CATI) system. The survey sample was drawn from the total noninstitutionalized adult populations (ages 18 years and over) residing in telephone-equipped dwelling units in the three target cities. The telephone survey followed a 10-attempt dialing protocol, in which up to 10 attempts were made unless a final disposition was obtained. A final disposition was attained when (1) the respondent completed the interview, (2) the telephone number was found to be invalid, (3) the record reached 10 attempts distributed among three different day parts, or (4) the respondent gave a final refusal. Experienced, supervised personnel conducted the interviews using CATI software (Computers for Marketing Corp, San Francisco, California) The CSQ study provided for a disproportionally allocated, stratified, random digit sample of telephone-equipped residential households in the targeted sites. Each of the three sites, New York, Baltimore, and San Juan, was sampled independently. Both English and Spanish versions of the CSQ were available and used at the preference of the interviewee.

Table 1 shows the key questions from the CSQ that constituted the primary dependent variables, ie, willingness to participate in cancer screenings. Question 7 is the single best gestalt item in the CSQ to assess overall willingness to participate without regard to specific type of cancer, while the subsections of Question 16 ask about willingness to participate in 10 site-specific types of cancer screenings. The influence of "who conducts" the cancer screening is the focus of Question 17 as its subquestions present the respondent with a choice of 10 different providers. Finally, Question 18 focuses on the influence of "what one is asked to do" in the cancer screening on willingness to participate as it describes 11 different activities. The primary independent variable of race/ethnicity was obtained using the standard two-part question, first asking about Hispanic status (Yes/No), then asking about race using the standard US categories of white, black/African-American, Asian/Pacific Islander, American Indian/Native American, and Other. Respondents who indicated that they are Hispanic further specified whether they are Puerto Rican or Mexican-American, or Other Hispanic. The variable of age was calculated from a "date of birth" variable on the CSQ. The level of education and level of income variables were collected via an ordinal listing of nine ascending categories of educational level and of 10 ascending categories of income level, each of which was then collapsed into three ascending categories for the demographic table and five ascending categories for the multivariable analyses. To acknowledge and account for cultural differences between the cities (above and beyond simple demographic differences), the variable of "city" was included as a separate covariate in all multivariable analyses.

Table 1. — Four Questions From the CSQ on the Likelihood of Participation in Cancer Screening Examinations

Response choices for all four questions are:
- very likely
- somewhat likely
- not quite sure
- somewhat unlikely
- very unlikely

Question 7. How likely are you to agree to have a cancer screening exam at the present time? Are you … (*read response choices above*)?

Question 16. How likely are you to have each of the following specific types of cancer screening exams? If the screening exams were for …(*give specific exam from list below*), would you be …(*read responses choices above*)?
- breast cancer
- colon cancer
- lung cancer
- stomach cancer
- liver cancer
- prostate cancer
- blood cancer
- skin cancer
- rectal cancer
- oral cancer

Question 17. Some people would feel differently depending on who was providing the cancer screening. I'm going to read you a list of people who might do a cancer screening. For each one, tell me how likely you would be to participate in a cancer screening exam. If the screening exam were provided by …(*give specific "who" from list below*), would you be …(*read response choices above*)?
- own doctor
- university medical school
- university dental school
- insurance company
- drug company
- own dentist
- nonprofit foundation
- university nursing school
- government
- tobacco company

Question 18. There are different types of cancer screening exams that involve different things. How likely are you to participate in a cancer screening exam if you had to do the following? How about if you had to …(*give specific activity from list below*), would you be (*read response choices above*):
- give blood sample
- exam by doctor
- spit out saliva
- scrape cells from mouth/skin
- exam by nurse
- surgical skin biopsy
- x-rays taken
- diet interview
- exam by dentist
- smoking habit interview
- alcohol habit interview

Unadjusted bivariable analysis was used as a pathway leading to adjusted multivariable analyses. For the unadjusted bivariable analyses, the proportion of those willing to participate in cancer screenings were dichotomized into Likely (Very Likely + Somewhat Likely) and Unlikely (Not Quite Sure + Somewhat Unlikely + Very Unlikely) and statistical significance was evaluated by means of chi-square tests with the significance level set at $P \leq .05$. Multivariable logistic regression analysis, which accounted for the multistage sampling techniques used in the random-digit-dial telephone survey and which adjusted for age, sex, education, income, and city, produced odds ratios (ORs) with 95% confidence intervals (CIs). Specifically, the PROC SURVEYFREQ and PROC SURVEYLOGISTIC procedures (SAS Institute Inc, Cary, North Carolina) were used for all data analyses to account for the complex sampling frame used in this random-digit-dial survey. Finally, the fit of the logistic regression was assessed by the omnibus test, the Hosmer-Lemeshow Test, as well as the classification table and the residual scores.

Results

The CSQ response rates (ie, number of completed interviews/number of contacted households) ranged from 45% to 58% by city, with an overall completion rate (ie, number of completed interviews/number of initiated interviews) of 82.6%. The original targeted within-city enrollment goals for each ethnic/racial group were met or exceeded, and the overall final CSQ study sample (N = 1,148) consisted of 355 blacks, 311 Puerto Rican Hispanics, and 482 non-Hispanic whites. The mean age of respondents was 46.8 years (SD = 16.0), with a range of 19 to 96 years, and 65.1% of the respondent sample were female. Table 2 shows the age, sex, education, and income distribution of the 1,148 subjects within the three racial/ethnic groups.

In response to the question: "How likely are you to agree to have a cancer screening exam at the present time?" (Question 7), considered to be the overall single

best gestalt question on willingness to participate in cancer screenings, the vast majority of each racial/ethnic group indicated that they were likely to participate in cancer screenings: 79.5% for Puerto Rican Hispanics, 73.7% for blacks, and 73.2% for whites (differences were not statistically significant, $P = .11$). Even after a multivariable logistic regression analysis adjusted for age, sex, education, income, and city, there were no statistically significant differences in willingness to participate in cancer screenings across the three racial/ethnic groups, as measured by Question 7.

The differences by race/ethnicity in percent willing to participate in cancer screenings for different site-specific type-of-cancer examinations (Question 16a–j) are shown in Fig 1. The findings, for all subjects combined, ranged from a high in self-reported willingness to participate for the two sex-specific cancers (93% for breast cancer examinations and 85% for prostate cancer examinations) to a low for oral cancer examinations (70%). For the remaining seven site-specific type-of-cancer screening examinations, the findings revealed that the self-reported willingness to participate ranged from 71% to 77%. (The percent of "willingness to participate for skin cancer" in Fig 1 was calculated as 74% by using only the average "willingness to participate" percent of whites and Hispanics combined without including the

Table 2. — Distribution of Age, Sex, Education, and Income Within Racial/Ethnic Groups in the 3-City Cancer Screening Study (N = 1,148)

Race/Ethnic Group	Mean Age (± SD)	Female	Education Level	Income Level
Blacks (n = 355)[a,b]	45.1 ± 16.5	72.4%	Less than high school graduate = 19.9% High school graduate = 59.1% College graduate or greater = 21.0%	< $20,000 = 41.5% $20,000–$74,999 = 49.5% ≥ $75,000 = 9.0%
Whites (n = 482)[a,c]	47.5 ± 17.0	56.8%	Less than high school graduate = 10.6% High school graduate = 44.0% College graduate or greater = 45.4%	< $20,000 = 19.6% $20,000–$74,999 = 59.7% ≥ $75,000 = 20.6%
Puerto Rican Hispanics (n = 311)[b,c]	44.0 ± 15.9	69.5%	Less than high school graduate = 19.8% High school graduate = 43.3% College graduate or greater = 36.9%	< $20,000 = 39.2% $20,000–$74,999 = 49.8% ≥ $75,000 = 11.0%

Statistically significant contrasts*:
[a] For blacks vs whites contrast: differed on age, sex, education and income ($P \leq .05$).
[b] For blacks vs Puerto Rican Hispanics contrast: differed on education ($P \leq .05$).
[c] For Puerto Rican Hispanics vs whites contrast: differed on age, sex, education and income ($P \leq .05$).
* = ANOVA test for age; Pearson chi-square tests for gender, education, and income.

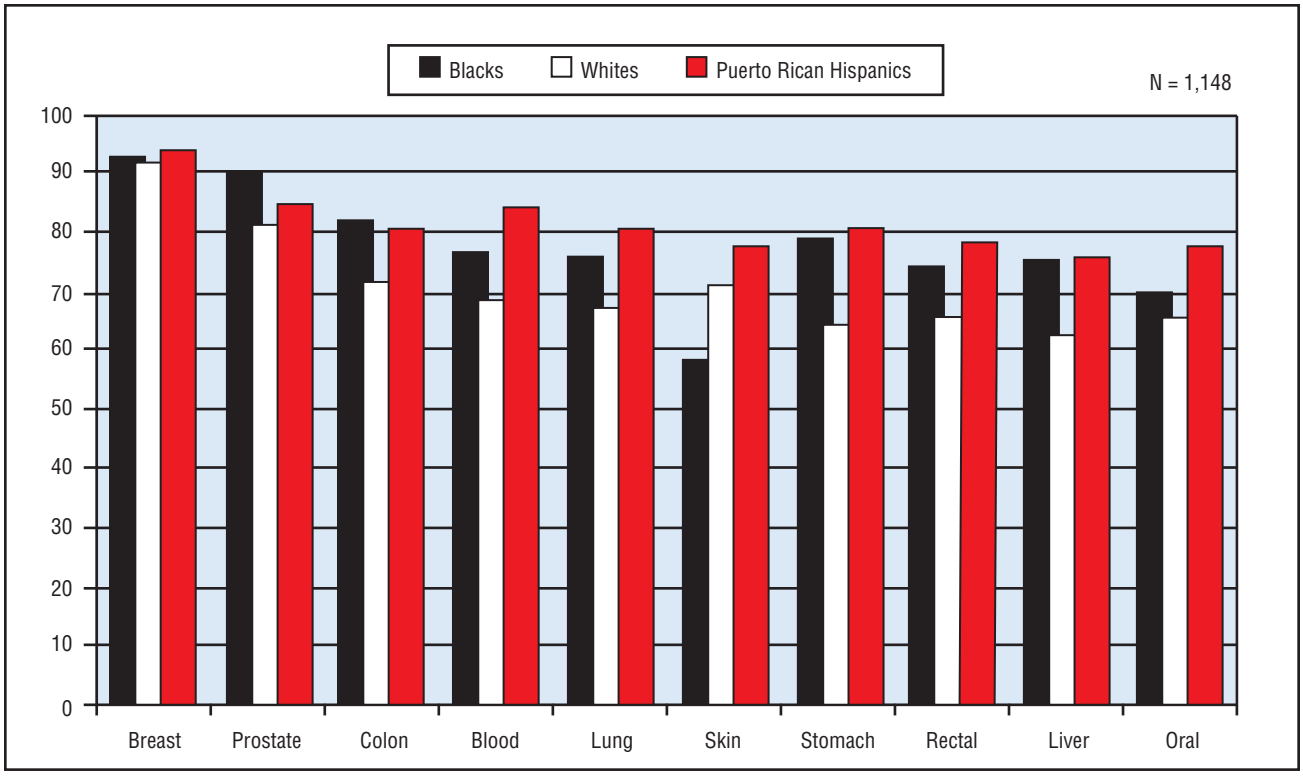


Fig 1. — Percentage likely to participate in cancer screenings dependent on type of cancer being screened for in the 3-City Cancer Screening Study.

percent for blacks, to acknowledge the reality of much lower skin cancer risk in blacks and thus to avoid biasing the "best" estimate for this specific cancer.)

The observed racial/ethnic differences in willingness to participate in cancer screening examinations for the eight non-sex-specific cancer sites were all statistically significant ($P < .002$) with a clear pattern among the three racial ethnic groups. Whites were consistently the least likely to indicate willingness to participate, and Puerto Rican Hispanics were consistently the most willing. The only exceptions were colon cancer screenings (for which blacks were the most willing) and skin cancer screenings (with blacks appropriately seeing themselves at lowest risk and thus least willing). While there were no statistically significant differences between the racial/ethnic groups for the two sex-specific cancers (breast and prostate) regarding willingness to participate in cancer screenings, the observed difference for prostate cancer examinations was notable with blacks (91%) being more willing than whites (81%) ($P = .096$). This percent difference likely failed to achieve statistical significance due to the smaller sample size, ie, only male subjects or 35% of the total sample size for the non-sex-specific cancers.

As shown in Table 3, when a multivariable logistic regression analysis (adjusted for age, sex, education, income, and city) was performed for each of the 10 site-specific types of cancer, the ORs indicated that (1) Puerto Rican Hispanics, compared with whites, were more willing to participate in cancer for 4 of the 10 site-specific cancers: lung, stomach, liver, and blood (leukemia) cancers (OR = 2.07–2.60), (2) blacks, compared with whites, were more willing to participate in a cancer screening for stomach cancer (OR = 2.13, 95% CI 1.24–3.65), and (3) blacks, compared with whites, were less willing to participate in skin cancer screenings (OR = 0.53, 95% CI 0.31–0.89). Except for this finding on skin cancer in blacks, neither blacks nor Puerto Rican Hispanics self-reported a lower willingness than whites for any other of the site-specific types of cancer.

The unadjusted bivariable findings from Question 17a–j are shown in Fig 2. These findings reveal, for the study sample as a whole, a large range in percent likely to participate in cancer screenings depending on "who" was conducting the study, from a high of 91% "if conducted by your own MD" to a low of 25% "if run by a tobacco company" (a 3.6-fold difference). Statistically significant differences across racial/ethnic groups were detected for five specific "who" categories (university dental school, government, insurance company, drug company, and tobacco company) with a consistent pattern across the three racial/ethnic groups. Puerto Rican Hispanics were the most willing to indicate that they would participate; whites were the least willing. The three racial/ethnic groups showed slight to occasionally substantial differences in response to any one "who" probe, but on the whole they exhibited similar ratings across the "who" factor (Fig 2). The three racial/ethnic groups ranked "own doctor," "own dentist," and "university medical school" as the most trusted in "who" factors, while "drug company" and "tobacco company" ranked lowest. Among the other "who" factors that formed the middle responses (about 45% to 58% likelihood), "government" was the lowest of these middle-ranked factors (45%).

Table 3 also shows the multivariable logistic regression analysis (adjusted for age, sex, education, income, and city) performed for the 10 different health care

Table 3. — Logistic Regression Multivariable Analyses for Statistically Significant Findings From the Key Questions From the CSQ by Racial/Ethnicity Group in the 3-City CSQ Study (N = 1,148)*

Question 16	**Specific Type of Cancer**	**Contrast**	**OR**	**95% CI**
	Skin	Black vs white	0.53	0.41–0.89
	Lung	Puerto Rican Hispanic vs white	2.16	1.20–3.87
	Stomach	Black vs white	2.13	1.24–3.65
	Stomach	Puerto Rican Hispanic vs white	2.60	1.43–4.70
	Liver	Puerto Rican Hispanic vs white	2.38	1.33–4.24
	Blood	Puerto Rican Hispanic vs white	2.07	1.06–4.05
Question 17	**"Who" Conducts Cancer Screening**	**Contrast**	**OR**	**95% CI**
	Nonprofit foundation	Black vs white	1.90	1.07–3.36
	Nonprofit foundation	Puerto Rican Hispanic vs white	1.96	1.07–2.58
	Tobacco company	Black vs white	1.78	1.08–2.92
	Drug company	Black vs white	2.06	1.28–3.31
	Drug company	Puerto Rican Hispanic vs white	2.64	1.54–4.52
Question 18	**"What Asked To Do" in Cancer Screening**	**Contrast**	**OR**	**95% CI**
	Exam by nurse	Black vs white	2.14	1.29–3.55
	Diet interview	Black vs white	2.97	1.44–6.11
	Alcohol habit interview	Puerto Rican Hispanic vs white	0.54	0.30–0.97
	Skin biopsy	Black vs white	0.45	0.28–0.72

* Adjusted for race, age, sex, education, income, and city, and accounting for multistage sampling. OR = odds ratio, CI = confidence interval.

providers or health agencies/institutions (ie, focused on the factor of "who" conducted the cancer screenings factor). The only statistically significant finding revealed ORs for blacks and Puerto Rican Hispanics that were higher than whites for willingness to participate when the "who" was either a nonprofit foundation, drug company, or tobacco company (the latter only for blacks). ORs ranged from 1.78 to 2.64. A check of the magnitude of the regression coefficients in these analyses revealed a good fit of the logistic regression model as used. Further in this exploratory first use of the CSQ in a survey, as none of the nonstatistically significant findings showed any noteworthy differences, they were not judged to be meaningful.

In parallel fashion, the unadjusted bivariable findings from Question 18a–k on "what one is asked to do

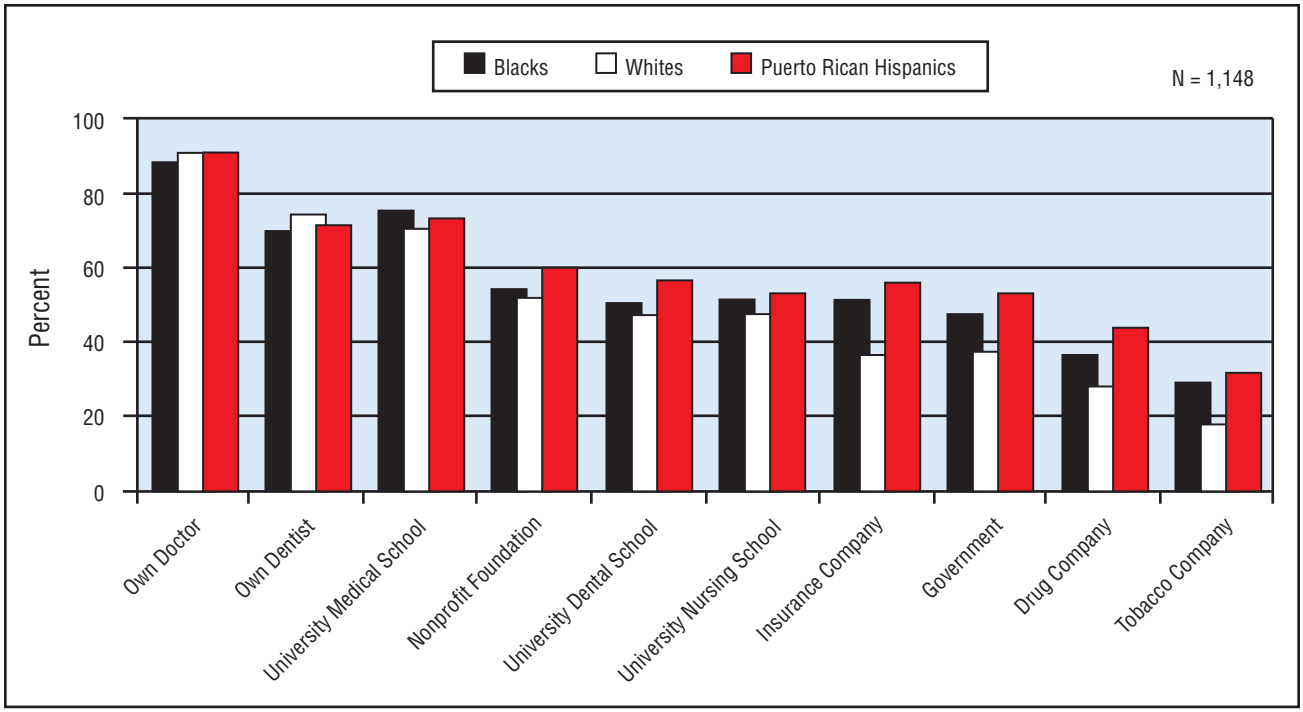


Fig 2. — Percentage likely to participate in cancer screenings dependent on "who" was conducting the screening in the 3-City Cancer Screening Study.

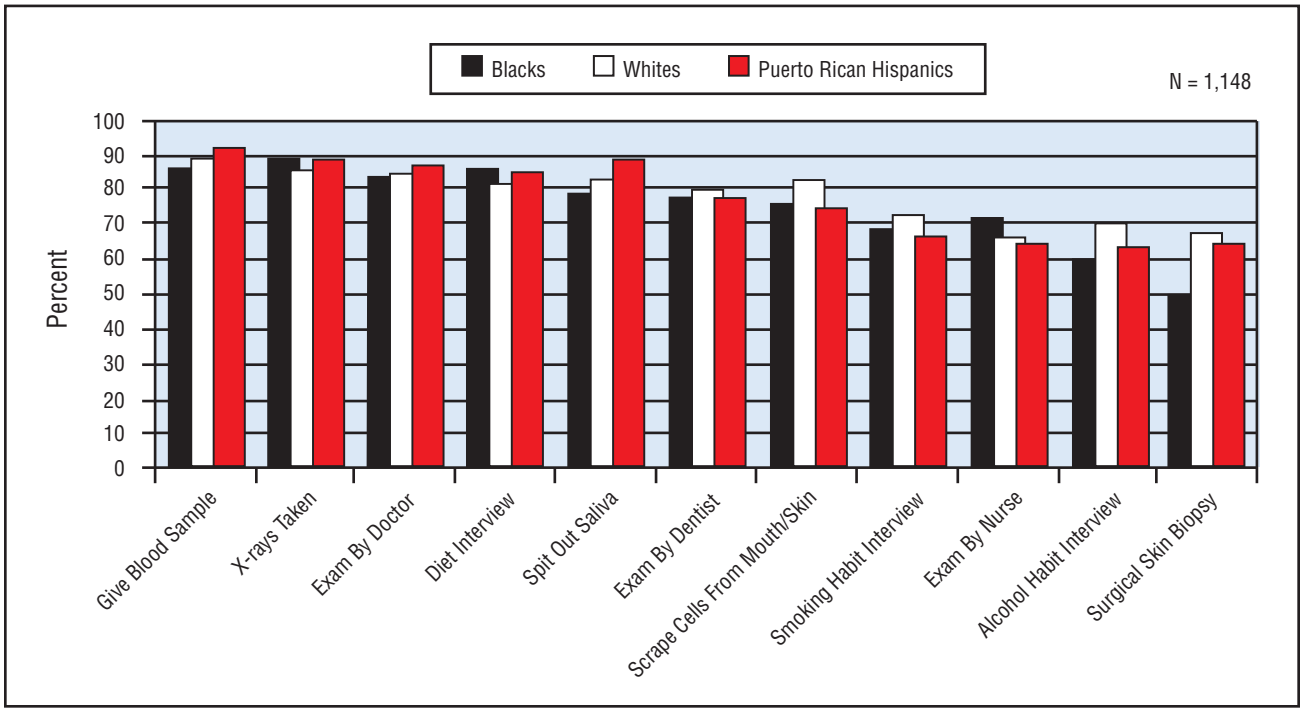


Fig 3. — Percentage likely to participate in cancer screenings dependent on "what one is asked to do" in the screening in the 3-City Cancer Screening Study.

in a study" are presented in Fig 3. Again, the range is considerable, depending on "what one is asked to do" in cancer screenings, and the three racial/ethnic groups generally demonstrated similar ratings across the 11 specific probes; they appeared to more or less travel together "up-and-down" the scale of willingness to participate, with no consistent pattern by race/ethnicity. The "what one is asked to do" categories that elicited the highest willingness to participate were "give blood sample" and "have x-rays taken" (89% and 88%, respectively) followed by "exam by doctor" (85%) and "diet interview" (84%). The middle group of categories were composed of "spit out saliva" (80%), "exam by dentist" (79%) and "scrape cells from mouth or skin" (79%). The lowest group included "smoking habit interview" (70%), "exam by nurse" (68%), "alcohol habit interview" (66%), with "surgical skin biopsy" having the lowest ranking (61%). Statistically significant differences between the racial/ethnic groups was found for only four of these "what one is asked to" categories: whites were highest and blacks lowest for "alcohol habit interviews" and "surgical skin biopsy," whites were highest and Puerto Rican Hispanics lowest for "scraping cells from mouth or skin," and Puerto Rican Hispanics were highest and blacks lowest for "give blood sample."

Multivariable logistic regression analysis (adjusted for age, sex, education, income, and city) performed for the 11 differing categories of "what one is asked to do" in the cancer screenings activities (Table 3) revealed statistically significant findings. The OR for blacks, compared with whites, to self-report willingness to participate in a cancer screening was higher when they were asked to be interviewed about diet habits (OR = 2.97, 95% CI 1.44–6.11) or asked about having an examination by a nurse (OR = 2.14, 95% CI 1.29–3.55). However, the OR for blacks, compared with whites, to self-report willingness to participate in a cancer screening when they were asked to have a surgical skin biopsy taken under local anesthetic was less than one-half that of whites (OR = 0.45, 95% CI 0.28–.072).

Fig 4 shows the statistically significant differences between blacks, Puerto Rican Hispanics, and whites presented in Table 3 on the willingness to participate in cancer screenings for site-specific type of cancer, for "who" does the examination, and for "what one is asked to do" in that cancer screening for the specific ethnic/racial group contrasts.

Discussion

Given that the overall aim of the 3-City Cancer Screening Study was to address a range of issues related to factors that determine whether blacks and Puerto Rican Hispanics, compared with non-Hispanics whites, would be willing to take part in cancer screening examinations, there is clear justification for presenting both the unadjusted bivariable analyses and the adjusted multi-

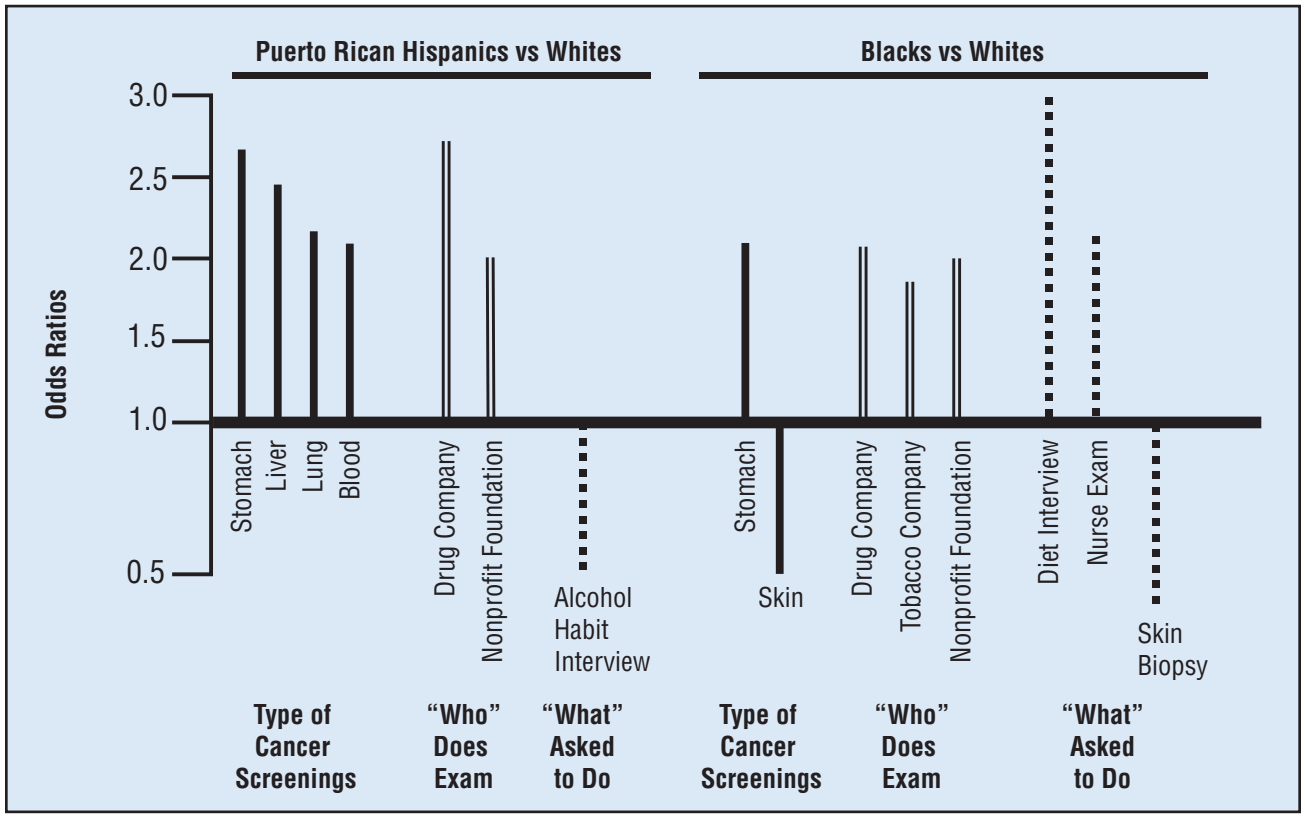


Fig 4. — Statistically significant odds ratios for likelihood of participating in cancer screenings: a comparison of blacks vs whites and of Puerto Rican Hispanics vs whites, type of cancer screenings, "who does exam," and "what" one is asked to do in the screening in the 3-City Cancer Screening Study.

variable analyses since each has specific utility to different audiences. The unadjusted bivariable analyses has utility for those in the public health arena since the data provided have immediate "on the ground" reality; these data regarding a racial/ethnic group currently would be valid. Basically, the unadjusted bivariable analyses target the "what happens if I plan to do this in the community" type of question. On the other hand, adjusted multivariable analysis results possess a different utility, one more related to a long-term, more detached "controlled scientific understanding" of the factors that affect any observed differences across race/ethnicity. Specifically, multivariable analyses pursue the more analytical (albeit often hypothetical) "what if" type of question, seeking to search for the influence of a key factor of interest under the "artificial situation" where all other factors are kept equal across the racial/ethnic groups.

Thus, the answer to the question posed by public health workers — "Are minorities less likely to be willing to participate in cancer screenings than whites?" — is found in the unadjusted bivariable analyses, in the data that reflect all of the cultural and demographic imbalances in today's communities. The findings show that the overall answer is "no," based on these self-reported data on willingness to participate in cancer screenings. Across 62 separate statistical contrasts between blacks and whites or Hispanics and whites for the 31 different specific circumstances of "type of cancer" screenings, "who" conducts the screening, and "what one is asked to do" in the screening, blacks and/or Hispanics were no less likely than whites to self-report willingness to participate in 93.5% of these specific contrasts.

Based on the adjusted multivariable analyses, the key scientific findings on influential factors show that race/ethnicity does have a measurable influence on many of the observed statistically significant differences in self-reported willingness to participate in cancer screenings when age, sex, education, income, and city are taken into account. For 12 of the 15 statistically significant multivariable adjusted contrasts between racial/ethnic groups (across the 62 contrasts for "type of cancer," "who conducts the screening," and "what one is asked to do"), the odds of being willing to participate were 2- to 3-fold higher for blacks and/or Hispanics than for whites. There were three exceptions in which a minority group was less likely to self-report willingness to participate than whites. Two involved blacks and skin cancer circumstances (for skin cancer screening, and having a skin biopsy) and one involved Puerto Rican Hispanics (having a alcohol habit interview).

A limitation of this study was that it was the first administration of the CSQ in a full survey, and therefore methodological questions remain to be answered about the CSQ as a research instrument. The CSQ was derived from the previously designed questionnaire, the Tuskegee Legacy Project (TLP) Questionnaire, which was designed to explore whether minorities are more reluctant than whites to participate in biomedical research, and if so, to explore the reasons (including the Tuskegee Syphilis Study) for any observed differences. The TLP Questionnaire was developed over a period of 3 years via focus group studies and a series of pilot studies and is described elsewhere.[31,32] Subsequently, the same multidisciplinary team that developed the TLP Questionnaire on "willingness to participate in biomedical research studies" then developed a parallel questionnaire, the CSQ, to explore similar issues regarding "willingness to participate in cancer screenings." As the CSQ was derived in format from the TLP Questionnaire, most questions were already refined via the focus group and pilot studies were done in developing the TLP Questionnaire over that 3-year period. Necessary changes in wording to address "cancer screening participation" were pretested on a small sample of pilot CSQ interviews. The current wording of questions related to cancer screening participation as used in the CSQ reflects how invitations to have screenings for a specific type of cancer are typically announced in a community or referred to by physicians, and thus they likely have face validity, as was demonstrated in our pilot study phase. Nevertheless, future methodological studies could probe for a clearer understanding of the beliefs of the respondents when they answered questions about willingness to participate in different types of cancer screenings. Was it a general concern about that type of cancer, or were they well enough informed to envision different specific cancer screening procedures as used by physicians (eg, FOBT vs colonoscopy for colon cancer)? These and other in-depth methodological probes would serve to improve both the CSQ as a research instrument (via modifications) and thereby improve our understanding of individuals' willingness to participate in cancer screenings as a result of either additional questions developed or modifications to existing questions currently used in the CSQ.

The burden of cancer falls disproportionately on the poor and disadvantaged in society, but the reasons for these disproportionate cancer-related health disparities are still unclear. However, although there is consensus that early detection of cancer is one of the most effective means of assuring timely treatment and survival, too few people take advantage of the tests available to screen early for common cancers.[10,11] This is especially true for Hispanics and blacks who have lower screening rates than whites have.[5,6,33] Overall, blacks and Hispanics, as well as others with lower level of formal education, are less likely to have such examinations, possibly due to reduced access to medical care.[7-9,12] Several other factors have been associated

with low cancer screening utilization: low income level, low level of education, lack of insurance,[6,7,34-36] older age,[36-40] unemployment, poor housing, inadequate access to health care,[1,36,41,42] limited knowledge about cancer, poor knowledge and attitude toward the screening process, lack of a regular physician, language barriers, and competing demands.[7,12-14,40,43-45]

Studies conducted with different minority groups have found that in addition to the barriers already mentioned, other important factors hinder cancer screening utilization for these populations. These factors include cultural differences, racial bias, emotions and beliefs,[46,47] fear of cancer, embarrassment, acculturation,[48,49] and fatalistic beliefs.[37,50] Thus, it is not surprising that the burden of cancer deaths is particularly high among blacks and Hispanics compared with whites.

Several recent surveys have asked respondents about their past cancer screening activities and participation. Surveys on past cancer screening participation rates have reported a range of 46% to 80% for breast cancer in women,[5,7-9] 56% to 82% for cervical cancer,[5,6,8,13] 36% to 55% for colon cancer,[12,14] and 41% for prostate cancer.[5] While the range of these findings of reported past cancer screening participation (36% to 82%) is lower than those in this current study on rates of self-reported willingness to participate for 10 site-specific cancer screenings (70% to 94%), the fact that the two ranges do overlap is an indication that "cancer screening behavior" does approach cancer screening intentions. Clearly, the high end for both practices and intentions were found for those cancer screening programs that have received major attention in the media and have the biggest campaigns for participation, such as breast and cervical cancers in women.

Conclusions

The findings from this study show that blacks and Hispanics self-report a willingness to participate in cancer screenings that is at least equal to that of whites, and they are frequently more willing than whites, under several different circumstances. The only exception involved blacks and skin cancer, where the odds of willingness to participate in cancer screenings were, appropriately and understandably, half that of whites. Given these findings, efforts by researchers, medical personnel, and community health workers are needed to overcome the often-cited barriers faced by these minority populations. Encouraging the willingness of minorities to take part in screening programs would equalize participation rates in screening and thus improve outcomes as regards to cancer mortality.

Appreciation is expressed to Mr Piotr Brzoza, Ms Michelle M. Haghpanah, and Ms Anne X. Truong, three graduate students, for their assistance with preliminary analyses related to this manuscript.

Disclosures

The 3-City Cancer Screening Study was supported by grant U54 DE 14257 from the National Institute of Dental and Craniofacial Research/National Institutes of Health (NIDCR/NIH), the New York University Oral Cancer Research on Adolescent and Adult Health Promotion Center (an Oral Health Disparities Research Center).

No significant relationship exists between the authors and the companies/organizations whose products or services may be referenced in this article.

References

1. National Cancer Institute Cancer Health Disparities: Questions and Answers. http://www.cancer.gov/cancertopics/factsheet/cancer-health-disparities. Accessed May 21, 2008.
2. Centers for Disease Control and Prevention (CDC). Invasive cervical cancer among Hispanic and non-Hispanic women: United States, 1992-1999. *MMWR Morb Mortal Wkly Rep*. 2002;51(47):1067-1070.
3. Puerto Rico Central Cancer Registry, Preliminary data on death rates-2003, Puerto Rico Dept of Health, November 2006.
4. Puerto Rico Central Cancer Registry, Preliminary data on incidence rates-2003, Puerto Rico Dept of Health, November 2006.
5. Swan J, Breen N, Coates RJ, et al. Progress in cancer screening practices in the United States: results from the 2000 National Health Interview Survey. *Cancer*. 2003;97(6):1528-1540.
6. Zambrana RE, Breen N, Fox SA, et al. Use of cancer screening practices by Hispanic women: analyses by subgroup. *Prev Med*. 1999;29(6, pt 1): 466-477.
7. Garza MA, Luan J, Blinka M, et al. A culturally targeted intervention to promote breast cancer screening among low-income women in East Baltimore, Maryland. *Cancer Control*. 2005;12(suppl 2):34-41.
8. Coughlin SS, Uhler RJ. Breast and cervical cancer screening practices among Hispanic women in the United States and Puerto Rico, 1998-1999. *Prev Med*. 2002;34(2):242-251.
9. Coughlin SS, Uhler RJ, Bobo JK, et al. Breast cancer screening practices among women in the United States, 2000. *Cancer Causes Control*. 2004;15(2):159-170.
10. Jemal A, Siegel R, Ward E, et al. Cancer statistics, 2006. *CA Cancer J Clin*. 2006;56(2):106-130.
11. Smith RA, Cokkinides V, Eyre HJ. American Cancer Society guidelines for the early detection of cancer, 2006. *CA Cancer J Clin*. 2006;56(1): 11-25; quiz 49-50.
12. Wee CC, McCarthy EP, Phillips RS. Factors associated with colon cancer screening: the role of patient factors and physician counseling. *Prev Med*. 2005;41(1) 23-29. Epub 2004 Dec 29.
13. Abraído-Lanza AF, Chao MT, Gammon MD. Breast and cervical cancer screening among Latinas and non-Latina whites. *Am J Public Health*. 2004;94(8):1393-1398.
14. James AS, Campbell MK, Hudson MA. Perceived barriers and benefits to colon cancer screening among African Americans in North Carolina: how does perception relate to screening behavior? *Cancer Epidemiol Biomarkers Prev*. 2002;11(6):529-534.
15. Horowitz AM, Nourjah PA. Factors associated with having oral cancer examinations among US adults 40 years of age or older. *J Public Health Dent*. 1996;56(6):331-335.
16. Claudio C, Katz RV, Green BL, et al. Cancer screening participation: comparative willingness of San Juan Puerto Ricans versus New York City Puerto Ricans. *J Natl Med Assoc*. 2007;99(5):542-549.
17. Blocker DE, Romocki LS, Thomas KB, et al. Knowledge, beliefs and barriers associated with prostate cancer prevention and screening behaviors among African-American men. *J Natl Med Assoc*. 2006;98(8):1286-1295.
18. Forrester-Anderson IT. Prostate cancer screening perceptions, knowledge and behaviors among African American men: focus group findings. *J Health Care Poor Underserved*. 2005;16(4 suppl A):22-30.
19. McFall SL, Hamm RM, Volk RJ. Exploring beliefs about prostate cancer and early detection in men and women of three ethnic groups. *Patient Educ Couns*. 2006;61(1):109-116. Epub 2005 Oct 26.
20. Yu MY, Wu TY. Factors influencing mammography screening of Chinese American women. *J Obstet Gynecol Neonatal Nurs*. 2005;34(3):386-394.
21. Skokar NK, Vernon SW, Weller SC. Cancer and colorectal cancer: knowledge, beliefs, and screening performance of a diverse population. *Fam Med*. 2005;37(5):341-347.
22. Ogedegbe G, Cassells AN, Robinson CM, et al. Perceptions of barriers and facilitators of cancer early detection among low-income minority women in community health centers. *J Natl Med Assoc*. 2005;97(2):162-170.
23. Woods VD, Montgomery SB, Herring RP. Recruiting Black/African American men for research on prostate cancer prevention. *Cancer*. 2004; 100(5):1017-1025.
24. Jones AR, Thompson CJ, Oster RA, et al. Breast cancer knowledge, beliefs, and screening behaviors among low-income, elderly black women. *J Natl Med Assoc*. 2003;95(9):791-797, 802-805.

25. Gupta A, Kumar A, Stewart DE. Cervical cancer screening among South Asian women in Canada: the role of education and acculturation. *Health Care Women Int*. 2002;23(2):123-134.
26. McCoy CB, Anwyl RS, Metsch IR, et al. Prostate cancer in Florida: knowledge, attitudes, practices, and beliefs. *Cancer Pract*. 1995;3(2):88-93.
27. Morgan C, Park E, Cortes DE. Beliefs, knowledge, and behavior about cancer among urban Hispanic women. *J Natl Cancer Inst Monogr*. 1995;(18):57-63.
28. Tortolero-Luna G, Glober GA, Villarreal R, et al. Screening practices and knowledge, attitudes, and beliefs about cancer among Hispanic and non-Hispanic white women 35 years or older in Nueces County, Texas. *J Natl Cancer Inst Monogr*. 1995;(18):49-56.
29. Watts BG, Vernon SW, Myers RE, et al. Intention to be screened over time for colorectal cancer in male automotive workers. *Cancer Epidemiol Biomarkers Prev*. 2003;12(4):339-349.
30. Demark-Wahnefired W, Strigo T, Catoe K, et al. Knowledge, beliefs, and prior screening behavior among blacks and whites reporting for prostate cancer screening. *Urology*. 1995;46(3):346-351.
31. Katz RV, Kegeles SS, Green BL, et al. The Tuskegee Legacy Project: history, preliminary scientific findings and unanticipated societal benefits. *Dent Clin North Am*. 2003;47(1):1-19.
32. Katz RV, Kegeles SS, Kressin NR, et al. The Tuskegee Legacy Project: willingness of minorities to participate in biomedical research. *J Health Care Poor Underserved*. 2006;17(4):698-715.
33. Brown DR, Topcu M. Willingness to participate in clinical treatment research among older African Americans and Whites. *Gerontologist*. 2003; 43(1):62-72.
34. Palmer RC, Fernandez ME, Tortolero-Luna G, et al. Acculturation and mammography screening among Hispanic women living in a farmworker communities. *Cancer Control*. 2005;12(suppl 2):21-27.
35. Cancer Trends Progress Report: 2005 Update. National Cancer Institute, National Institute of Health, Department of Health & Human Services, Bethesda, MD; 2005. http://progressreport.cancer.gov/faq.asp?did=2005&fid=2005&mid=vfq&pid=1. Accessed May 28, 2008.
36. Bowie JV, Curbow BA, Garza MA, et al. A review of breast, cervical, and colorectal cancer screening interventions in older women. *Cancer Control*. 2005;12(suppl 2):58-69.
37. Mayo RM, Erwin DO, Spitler HD. Implications for breast and cervical cancer control for Latinas in the rural South: a review of the literature. *Cancer Control*. 2003;10(5):60-68.
38. Ramirez AG, Suarez L, Laufman S, et al. Hispanic women breast and cervical cancer knowledge, attitudes, and screening behaviors. *Am J Health Promot*. 2000;14(5):292-300.
39. Wu ZH, Black SA, Markides KS. Prevalence and associated factors of cancer screening: why are so many older Mexican-American women never screened? *Prev Med*. 2001;33(4):268-273.
40. Mandelblatt JS, Gold K, O'Malley AS, et al. Breast and cervix cancer screening among multiethnic women: role of age, health and sources of care. *Prev Med*. 1999;28(4):418-425.
41. Olsen S. Cancer prevention and screening for underserved populations. Adult Health. 1999;1(6). In HTML format accessed through Meniscus Educational Institute. http://www.meniscus.com/web/pages/homepage.asp.
42. Christian MC, Trimble EL. Increasing participation of physicians and patients from underrepresented racial and ethnic groups in National Cancer Institute-sponsored clinical trials. *Cancer Epidemiol Biomarkers Prev*. 2003; 12(3):277s-283s.
43. McCoy CB, Preyra M, Metsch LR, et al. A community-based breast cancer-screening program for medically underserved women: its effect on disease stage at diagnosis and on hazard of death. *Rev Panam Salud Publica*. 2004;15(3):160-167.
44. Luquis RR, Villanueva IJ. Knowledge, attitudes, and perceptions about breast cancer and breast cancer screening among Hispanic women residing in South Central Pennsylvania. *J Community Health*. 2006;31(1): 25-42.
45. Hannon P, Harris JR, Martin D, et al. Colorectal cancer screening in Washington State: predictors of current screening and explanations for no screening. *Prev Chronic Dis*. 2005;2.
46. Weissman JS, Schneider EC. Social disparities in cancer: lessons from a multidisciplinary workshop. *Cancer Causes Control*. 2005;16(1):71-74.
47. Magai C, Consedine N, Conway F, et al. Diversity matters: unique populations of women and breast cancer screening. *Cancer*. 2004;100(11): 2300-2307.
48. Wong-Kim E, Sun A, DeMattos MC. Assessing cancer beliefs in a Chinese immigrant community. *Cancer Control*. 2003;10(5 suppl):22-28.
49. O'Malley AS, Kemer J, Johnson AE, et al. Acculturation and breast cancer screening among Hispanic women in New York City. *Am J Public Health*. 1999;89(2):219-227.
50. Chavez LR, Hubbell FA, Mishra SI, et al. The influence of fatalism on self-reported use of Papanicolaou smears. *Am J Prev Med*. 1997;13(2): 418-424.

The Legacy of the Tuskegee Syphilis Study: Assessing its Impact on Willingness to Participate in Biomedical Studies

Ralph V. Katz, DMD, MPH, PhD
B. Lee Green, PhD
Nancy R. Kressin, PhD
S. Stephen Kegeles, PhD
Min Qi Wang, PhD
Sherman A. James, PhD
Stefanie L. Russell, DDS, MPH, PhD
Cristina Claudio, PhD
Jan M. McCallum, MSPH

Abstract: The phrase, *legacy of the Tuskegee Syphilis Study*, is sometimes used to denote the belief that Blacks are more reluctant than Whites to participate in biomedical research studies because of the infamous study of syphilis in men run by the U.S. Public Health Service from 1932–72. This paper is the first to attempt to assess directly the accuracy of this belief within a multi-city, multi-racial, large-scale, detailed random survey. We administered the Tuskegee Legacy Project (TLP) Questionnaire to 826 Blacks and non-Hispanic White adults in three U.S. cities. While Blacks had higher levels of general awareness of the Tuskegee Syphilis Study, there was no association between either awareness or detailed knowledge of the Tuskegee Syphilis Study and willingness to participate in biomedical research, either

RALPH KATZ *is Professor and Chair of the Department of Epidemiology & Health Promotion at the New York University College of Dentistry in New York, where* ***STEFANIE RUSSELL*** *is an Assistant Professor.* ***B. LEE GREEN*** *is the Executive Director of the Office of Institutional Diversity at the H. Lee Moffitt Cancer Center & Research Institute in Tampa, FL.* ***NANCY KRESSIN*** *is an Associate Professor in the Department of General Internal Medicine at the Boston University Medical School and a senior researcher at the Center for Health Quality, Outcomes & Economic Research Center in the Department of Veterans Affairs in Boston, MA.* ***S. STEPHEN KEGELES*** *(deceased) was a Professor in the Department of Behavioral Sciences and Community Health at the University of Connecticut School of Dental Medicine in Farmington, CT.* ***MIN QI WANG*** *is a Professor in the Department of Public and Community Health at the University of Maryland School of Public Health in College Park, MD.* ***SHERMAN JAMES*** *is the Susan B. King Professor of Public Policy in the in the Terry Sanford Institute of Public Policy and the Department of Community Medicine and Family Medicine at Duke University in Durham, North Carolina.* ***CRISTINA CLAUDIO*** *is a Professor in the Department of Health Behavior at the University of Puerto Rico School of Dentistry in San Juan, Puerto Rico.* ***JAN MCCALLUM*** *is a doctoral graduate student and Pre-doctoral Fellow in the Center for Health Promotion and Prevention Research at the University of Texas Health Science Center at Houston in Houston, Texas. Please address correspondence to Dr. Ralph Katz, Department of Epidemiology & Health Promotion, NYU College of Dentistry, 250 Park Ave. South, New York, NY 10003-1402; (212) 998-9550; fax (212) 995-4436; ralph.katz@nyu.edu.*

Journal of Health Care for the Poor and Underserved 19 (2008): 1168–1180.

for Blacks or Whites observed in our survey. While this study refutes the notion that there is a direct connection between detailed knowledge of the Tuskegee Syphilis Study and willingness to participate in biomedical research, it does not assess the broader question of whether and how historical events influence people's willingness to participate in research. Future studies should explore this.

Key words: Legacy of the Tuskegee Syphilis Study, minority participation in research, bioethics, health disparities.

There is widespread belief that the U.S. Public Health Service (USPHS) Syphilis Study at Tuskegee (1932–72) has had a negative effect on African-Americans' willingness to participate as research subjects in biomedical studies.[1–5] That study, long-recognized as unethical, committed abuses against 399 African-American sharecroppers in Macon Country, Alabama who were the subjects of a 40-year USPHS study of the effects of untreated syphilis in Black men.[6,7] While a considerable amount has been written about the long-lasting effects of the USPHS Syphilis Study at Tuskegee on the Black community, most of this work has been from the perspective of history,[8–15] ethics,[8–17] access to health care,[11,14–21] and law.[15,20]

Given the federal government's 1994 directive for researchers to obtain study samples with diversity in race and gender,[22] surprisingly little research has directly examined whether any differential participation of Blacks or other minorities in biomedical studies compared with Whites could be attributed to the USPHS Syphilis Study at Tuskegee,[23–30] and only three of these studies used quantitative methods.[28–30] A 2006 literature review pointed out the research design limitations of these three published quantitative studies on this specific topic[31] (they had low response rates,[30] were limited in scope to a single city,[28–30] or used study questionnaires containing only a limited set of variables[28–30]). Thus, while all three of these early exploratory studies concluded that awareness of the Tuskegee Syphilis Study did not affect willingness to participate in biomedical studies, they were limited in their scopes of inquiry as well as in their generalizability.

In our first paper from the Tuskegee Legacy Project data, we reported that, despite having greater fear of participation in research studies, Blacks were just as likely as Whites to self-report willingness to participate in biomedical research.[32] These initial findings were fully confirmed in our second major study on this same topic using the TLP Questionnaire in three different cities.[33] A third recent paper expanded our inquiry to address the impact of both the general awareness of the Tuskegee Syphilis Study and the 1997 Presidential Apology about the Tuskegee Syphilis Study on willingness to participate in biomedical studies; it reported that Blacks were 2–3 times *more* likely than Whites to be willing to participate in biomedical studies despite having heard of the Tuskegee Syphilis Study or the Presidential Apology.[34]

This paper seeks to answer a related but different question: "Does awareness and/or detailed knowledge of the Tuskegee Syphilis Study influence one's willingness to participate in biomedical studies in 1999–2000?" Specifically, we examined whether either general awareness or detailed knowledge of the Tuskegee Syphilis Study influences willingness to participate in biomedical studies for Blacks and for Whites in different geographic areas.

Methods

The Tuskegee Legacy Project (TLP) Study administered the TLP Questionnaire via random-digit dial (RDD) telephone interviews to respondents aged 18 years and older in four U.S. cities: Tuskegee, Alabama; Birmingham, Alabama; Hartford, Connecticut; and San Antonio, Texas in 1999–2000. This report presents the findings for the Black and White respondents from the three cities that had a sufficient number of Black and White respondents to permit meaningful analysis for both racial groups (i.e., Tuskegee, Birmingham, and Hartford).

As this paper is a follow-up analysis to our first paper from our Tuskegee Legacy Project, we refer readers to the earlier work for a full description of the methods used in conducting the 1999–2000 4-City Tuskegee Legacy Project.[32] This study was approved by the IRBs of the University of Connecticut Health Center and New York University.

The TLP Questionnaire contains two validated scales, the Likelihood of Participation (LOP) Scale (which measured the willingness to participate) and the Guinea Pig Fear Factor (GPFF) Scale (which measured the fear or wariness about participation), both of which are described in detail in the earlier publication.[32] Scores for each scale ranged from 0–100, with the LOP Scale using 17 questions from the TLP Questionnaire to provide a nuanced measure of willingness to participate in biomedical studies, while the GPFF Scale uses 5 questions to determine the level of fear an individual has about participating in a biomedical study. Awareness of the Tuskegee Syphilis Study (TSS) was measured dichotomously (i.e., yes/no) according to participants' responses to two separate questions, spaced 7 questions apart, asking if they had "ever heard of the Tuskegee Syphilis Study."

Additionally, the TLP Questionnaire contains a set of seven questions on detailed knowledge of the Tuskegee Syphilis Study (TSS) that were utilized to create the TSS Facts & Myths Quiz score, a Tuskegee Syphilis Study (TSS) knowledge quiz score (lowest = 0, highest = 7). For the calculation of the TSS Facts & Myths Quiz score, *don't know/not sure* responses were treated as missing values. The selection of the seven items emerged from discussions among the research team members based upon their familiarity with the Tuskegee Syphilis Study history and literature, with a guiding criterion that only well-established facts associated with the Tuskegee Syphilis Study were to be used, i.e., questions that had face vailidty. The items on the TSS Facts & Myths Quiz were reviewed by an NIH Study Section (as part of an overall grant protocol review process) which had no negative comments about either the inclusion of any one of the 7 items, or the claim that they constituted a reasonable measure of detailed knowledge about the Tuskegee Syphilis Study. Similarly, the entire TLP Questionnaire was pretested in face-to-face interviews in 3 small pilot studies, as well as in a final RDD pilot study of 200 subjects. It was never challenged by any of the over 300 subjects in these pre-test pilot studies.

Initial unadjusted bivariate analyses that examined frequency distributions were followed by multivariate general linear model ANCOVA analyses, and multiple linear regression analyses, for dichotomous or continuous dependent variables, respectively. The multivariate analysis tested an overall model of each dependent variable (LOP and GPFF) for the primary independent variable of interest (either awareness or

Box 1.

THE 7 ITEMS ON DETAILED KNOWLEDGE OF THE TUSKEGEE SYPHILIS STUDY (TSS) THAT CONSTITUTED THE QUESTIONS TALLIED TO CREATE THE TSS FACTS & MYTH QUIZ SCORE

The following was asked of subjects who had already indicated that they had heard of the Tuskegee Syphilis Study.

Introduction:

I would like to know what specific facts you remember about the Tuskegee Syphilis Study. Would you tell me whether the following statements are True or False. If you don't know, please tell me that.

7 items on the TSS Facts &Myth Quiz:	Factually Correct Answers for the TSS Facts & Myth Quiz Score[a]
1) Black men and women were subjects in the study	False
2) The subjects were injected with syphilis	False
3) The nurse who recruited them was Black	TRUE
4) The study lasted 40 years	TRUE
5) The subjects were told they had syphilis	False
6) The study was run by U.S. government doctors	TRUE
7) The study ended when penicillin was discovered as a cure for syphilis	False

[a]Total score could range from 0–7, with one point for each factually correct answer given.

detailed knowledge) with ethnicity, age, gender, education, and income as covariates at four strata: total sample, by city, by race, and by race-within-city. When significant factors were detected, additional analyses were conducted to assess the direction and the magnitude of the difference. If the significant factor had more than two levels, a *post hoc* test of the adjusted means with Tukey criterion was conducted to control for Type I error. All analyses were conducted using either SPSS v14.0 software[35] or SAS v9.0 software.[36]

Results

In this report, we present the TLP Questionnaire findings for the 826 adult Blacks and non-Hispanic Whites in three cities (Tuskegee, Alabama; Birmingham, Alabama; and Hartford, Connecticut) which had overall response rates of 65%, 70%, and 49%, respectively. The overall completion rate (number of completed interviews/number of

initiated interviews) exceeded 90% in each city. Table 1 shows the age, sex, education, and income distribution of the 826 subjects by race for each of the three cities.

In addition, Table 1 shows the proportion of Blacks and Whites, by city, who had "ever heard of the Tuskegee Syphilis Study." Awareness of the TSS ranged from a high of 93.3% for Blacks in Tuskegee to a low of 41.9% for Whites in Hartford. Across cities differences in the proportion of respondents who had ever heard of the TSS were statistically significant (89.9% vs. 70.9% vs. 49.0%, in Tuskegee, Birmingham, and Hartford, respectively; $p<.05$). Additionally, within city analyses by race revealed that a statistically significant ($p<.05$) higher percentage of Blacks than of Whites, had heard of the TSS in Birmingham (77.7% for Blacks vs. 64.0% for Whites) and Hartford (62.5% for Blacks vs. 41.6% for Whites), but not in Tuskegee (93.3% for Blacks vs. 87.8% for Whites).

The TSS Facts & Myth Quiz was administered to the 68.4% of subjects ($n=565$) who indicated that they had heard of the TSS. Figure 1 illustrates, by race within each city, the proportion of these 565 respondents who answered *true* to each of the 7 items on the Tuskegee Syphilis Study (TSS) Facts & Myth Quiz, having given a valid response (i.e., either *true* or *false*). All *don't know/not sure* responses for this analysis were treated as missing data; typically, across the 7 questions, the *don't know/not sure* responses constituted about 20% of the responses, with a low of 15% up to a high of 55%. In Figure 1, true statements are indicated in capital letters along the abscissa axis and any observed statistically significant differences by race within city are marked with asterisks. There was great variation across the 7 items by race within each city, with answers of *true* ranging from a high of 85% for the statement that subjects "were injected with syphilis" (a false statement) to a low of 2% for the statement that the study was "run by U.S. government doctors" (a true statement). For the former statement (i.e., subjects were injected with syphilis) the findings were that more Blacks than Whites thought it was true, and these differences were statistically significant ($p<.05$) for both Tuskegee and Hartford, and approached statistical significance ($p=.068$) in Birmingham.

Figure 2 reinforces the finding illustrated in Figure 1 concerning the relatively high frequency of incorrect answers on the details of the TSS for the TSS Facts & Myths Quiz. As Figure 2 shows, of a maximum 7 potential correct answers, neither Blacks nor Whites in any of the 3 cities achieved a mean TSS Facts & Myths score greater than 2, and no statistically significant racial differences were observed (mean scores $\pm$ standard deviations for Blacks $= 1.6\pm1.4$, and for Whites 1.7 ± 1.3). For the TSS Facts & Myth Quiz, 90% of the 565 respondents who had heard of the TSS had a detailed knowledge score of 3 or less, with the top scores achieved being 5 and 6 for Blacks and Whites, respectively.

Table 2 shows the findings from the multivariate analyses that were conducted to determine the impact of both awareness and detailed knowledge of the Tuskegee Syphilis Study on two dependent variables: 1) the willingness to participate (as measured by the LOP Scale), and 2) the degree of fear of participation (as measured by the GPFF Scale). Awareness of the Tuskegee Syphilis Study did not have a statistically significant relationship with either the willingness to participate (as measured by the LOP Scale, $p=0.71$) or the fear of participation (as measured by the GPFF Scale, $p=0.53$) when adjusted for race, age, sex, education, and income for the sample of 826 respondents

Table 1.

DEMOGRAPHIC CHARACTERISTICS BY RACE AND AWARENESS OF THE TUSKEGEE SYPHILIS STUDY (TSS) FOR TUSKEGEE, AL, BIRMINGHAM, AL AND HARTFORD, CT (N=826)

	Tuskegee, AL		Birmingham, AL		Hartford, CT	
	Black (n=104)	**White (n=180)**	**Black (n=103)**	**White (n=100)**	**Black (n=120)**	**White (n=219)**
Characteristic						
Age[a], years (± s.d.)	49.8 (±16.7)	56.1 (±17.0)	49.2 (±15.8)	51.7 (±16.8)	48.2 (±16.8)	52.7 (±16.0)
% male[b]	50.9	53.3	39.4	31.4	63.0	58.1
Education level[c], %						
< H.S	26.7	24.0	19.2	8.6	19.7	6.3
H.S. graduate	53.3	54.6	61.5	60.0	65.4	45.7
College graduate +	20.0	21.3	19.2	31.4	15.0	48.0
Income level[d]						
<$20,000	46.9	29.4	52.0	30.2	31.4	11.2
$20,000–74,999	51.0	63.8	44.0	55.2	60.2	51.6
>$74,999	2.0	6.7	4.0	14.6	8.5	37.2
Awareness of TSS[e] (% of respondents)	93.3%	87.8%	77.7%	60.0%	62.5%	41.9%

Note: Statistically significant findings ($p<.05$) across cities, and for Blacks vs. Whites within each city:

[a]**Age:** *within city finding*: Whites were older than Blacks in each city.

[b]**% male:** *across city finding*: Hartford had higher % male respondents than Tuskegee, which had higher % male respondents than Birmingham.

[c]**Education Level:** *across city finding*: Hartford had higher education levels than Birmingham, which had a higher education level than Tuskegee.

within city finding: Whites had higher educational levels in Hartford and in Birmingham than Blacks.

[d]**Income Level:** *across city finding*: Hartford had higher incomes than Birmingham and Tuskegee.

within city finding: Whites had higher incomes than Blacks in each city.

[e]**Awareness (%) of TSS:** *across city finding*: Respondents in Tuskegee had a higher awareness of the TSS than in Birmingham which had a higher awareness among respondents than in Hartford.

within city finding: A higher percentage of Blacks were aware of the TSS in Hartford and in Birmingham.

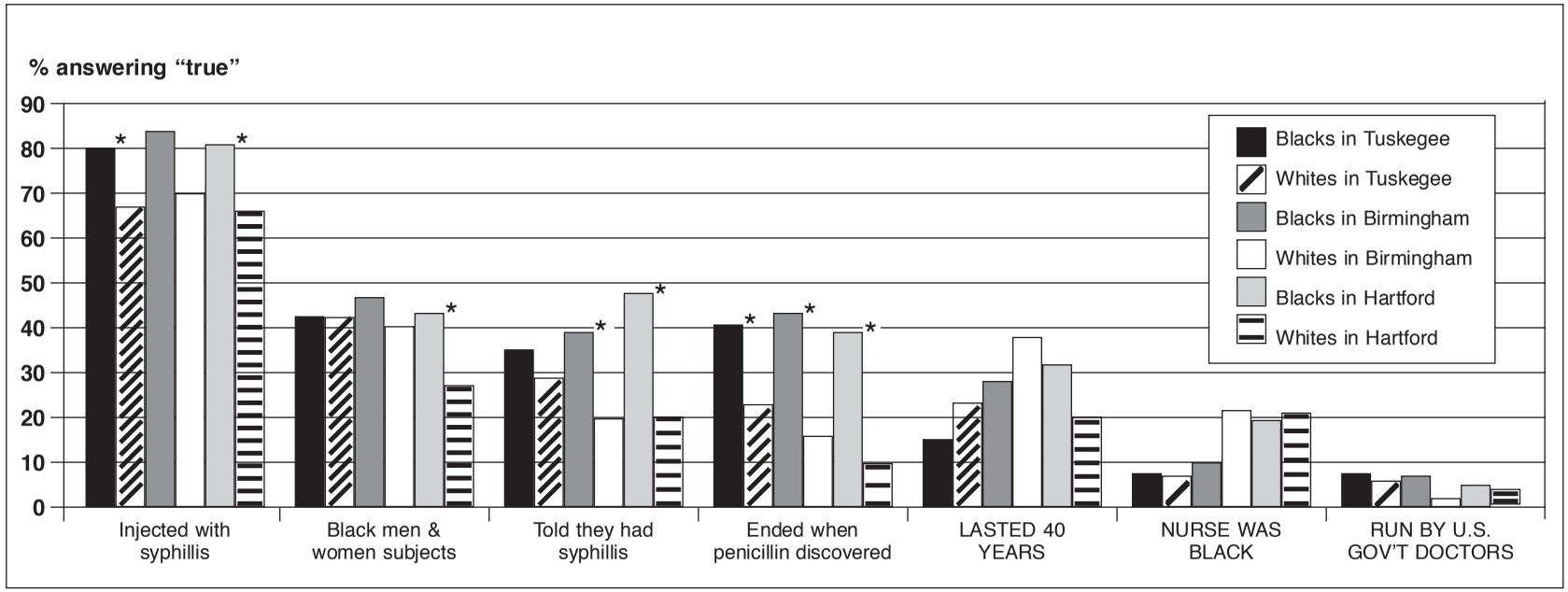


Figure 1. The percentage of respondents who answered 'true' to the 7 items on the Tuskegee Syphilis Study (TSS) Facts & Myth Quiz by race within city among those who had heard of the TSS (n=565).

Note: Capital letters indicate a TRUE statement.
*Statistically significant between Blacks vs. Whites within city at $p<.05$.

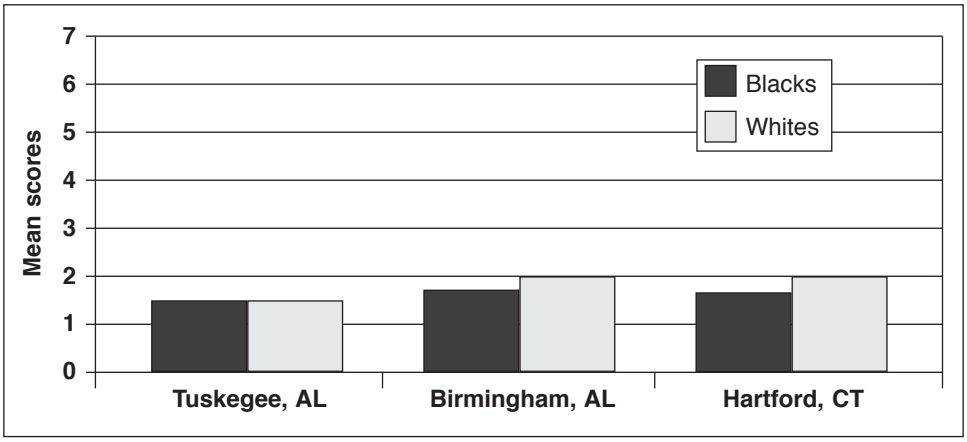


Figure 2. Mean scores* on TSS Facts & Myths Quiz by race within city for the 565 respondents who had heard of the Tuskegee Syphilis Study.

Note: none of the Black vs. Whites contrasts for either quiz score within city are statistically significant.
*(Mean # of correct answers).

taken as a whole, or for Whites when analyzed separately. However, an ANCOVA analysis by race of the relationship between awareness and GPFF in Blacks was statistically significant ($p=0.02$), with Blacks who were aware of the TSS having a higher mean GPFF score than Blacks who were not aware of the TSS (65.1 vs 57.3, $p=.02$). For the other ANCOVA analyses related to awareness by city and by race-within-city, no consistent pattern of significant findings was observed.

Table 2 also shows similarly adjusted ANCOVA analyses on the relationship between detailed knowledge of the TSS and the LOP and GPFF Scales. These analyses were conducted only for that subset of 565 respondents who indicated that they had heard of the TSS. The results were that detailed knowledge of the TSS, as measured by the mean Facts & Myth Quiz score, had no statistically significant relationship with either the LOP or GPFF Scale for Blacks or for Whites. However, a weak, albeit statistically significant inverse relationship was observed between the detailed knowledge and the GPFF Scale ($p=0.04$) for all subjects combined.

Discussion

The Tuskegee Legacy Project (TLP) had its origins in a 1994 bioethics conference held at the University of Virginia entitled *The Tuskegee Legacy: Doing Bad in the Name of Good*.[37] All the speakers at that conference discussed consequences of the Tuskegee Syphilis Study, especially the possibility that African Americans were more reluctant to participate in biomedical research because of the abuse they suffered in that infamous study. However, often speakers appeared to assume that the central hypothesis was true, i.e., that African Americans were, in fact, more reluctant to participate in biomedical studies and that the Tuskegee Syphilis Study was at the heart of this reluctance to become

Table 2.

SUMMARY OF ADJUSTED MULTIVARIATE ANALYSES[a] FINDINGS ON RELATIONSHIP BETWEEN AWARENESS AND DETAILED KNOWLEDGE OF THE TUSKEGEE SYPHILIS STUDY ON LIKELIHOOD OF PARTICIPATION (LOP) SCALE AND GUINEA PIG FEAR FACTOR (GPFF) SCALE FOR ALL SUBJECTS, WITHIN EACH CITY, AND WITHIN EACH CITY BY RACE

	Awareness[b] and		Detailed knowledge[c] and	
	LOP	GPFF	LOP	GPFF
All subjects	ns	ns	ns	p=.04[i]
By race:				
Blacks	ns	p=.02[g]	ns	ns
Whites	ns	ns	ns	ns
By city:				
Tuskegee	ns	ns	ns	ns
Birmingham	ns	ns	ns	ns
Hartford	p=.02[d]	ns	ns	p=.02[j]
By race within city:				
Tuskegee:				
Blacks	p=.02[f]	ns	ns	ns
Whites	ns	ns	ns	ns
Birmingham:				
Blacks	ns	ns	ns	ns
Whites	ns	ns	ns	ns
Hartford:				
Blacks	ns	ns	ns	p=.01[k]
Whites	p=.03[e]	ns	p=.02[h]	ns

[a]Adjusted for race, age, sex, education level, and income level with p=.05.
[b]Conducted for all 826 subjects in ANCOVA models.
[c]Conducted only for the 565 respondents who had heard of the TSS, using the TSS Facts & Myth Quiz as the independent variable in multiple linear regression models.
[d]In Hartford, those who were aware of the TSS had higher mean LOP scores compared with those who were not aware of the TSS (47.3 vs 41.9).
[e]Whites in Hartford who were aware of the TSS had higher mean LOP scores compared with Whites who were not aware of the TSS (49.7 vs 42.7).
[f]Blacks in Tuskegee who were aware of the TSS had lower mean LOP scores compared with Blacks who were not aware of the TSS (34.7 vs 55.9).
[g]Blacks, as a total group, who were aware of the TSS had higher mean GPFF scores compared with Blacks who were not aware of the TSS (65.1 vs 57.3}.
[h]Whites in Hartford who were aware of the TSS had an inverse relationship between their Detailed Knowledge score and LOP score (regression coefficient = −3.77).
[i]For all who were aware of the TSS, there was an inverse relationship between their Detailed Knowledge score and GPFF score (regression coefficient = −1.72).
[j]In Hartford, those who were aware of the TSS had an inverse relationship between their Detailed Knowledge score and GPFF score (regression coefficient = −3.35).
[k]Blacks in Hartford who were aware of the TSS had an inverse relationship between their Detailed Knowledge score and GPFF score (regression coefficient = −5.40).

a research subject. To date, most notions regarding this purported consequence of the Tuskegee Syphilis Study have been based on anecdote and, speculation, rather than on systematic empirical research.[5,9,11,20,23–26,29,30,37]

Our findings reveal a clear gradient across the three cities on the percentage of respondents who had ever heard of the Tuskegee Syphilis Study: nearly 90% having heard of the TSS in Tuskegee, the historical epicenter of the TSS; just over 70% in Birmingham, the nearest big city to Tuskegee with a university medical center that would conduct clinical studies; and just under 50% in Hartford, a distant city in a different region of the U.S. with a university medical center that would conduct clinical studies and with a demographic profile similar to Birmingham's. Our findings also showed that a lack of knowledge about the details of the TSS was fairly common among both Blacks and Whites (i.e., 90% of respondents gave factually correct answers on 3 or fewer of the 7 tested items).

The answer to our first central research question, "Did awareness of the Tuskegee Syphilis Study influence one's willingness to participate in biomedical studies in 1999–2000?" was a straightforward *No*, for both Blacks and for Whites. Our adjusted multilvariate analyses revealed no statistically significant relationships between awareness of the TSS and either the LOP or the GPFF Scales.

However, the findings within the city of Tuskegee differed from the findings in the cities of Birmingham and Hartford. A significant relationship was observed among Blacks between awareness of the TSS and willingness to participate, but only in the small city of Tuskegee, the epicenter of this infamous event.

The answer to our second central research question in this report, namely, "Among those who had heard of the Tuskegee Syphilis Study, did detailed knowledge of the Tuskegee Syphilis Study influence one's willingness to participate in biomedical studies in 1999–2000?" is also a straightforward *No* for Blacks and for Whites, when analyzed either together as a whole sample or separately. However, the data on fear of participation, as measured by the GPFF Scale for the sample as a whole (i.e., Blacks and Whites together), revealed a statistically significant but weak inverse relationship between GPFF and detailed knowledge for the study sample as a whole (but this finding did not hold for Blacks or for Whites when analyzed separately). This observation for the study sample as a whole regarding detailed knowledge may well be a spurious finding as the direction of the relationship is opposite to what one might anticipate, and thus may be an artifact simply of the large number (n=48) of statistical tests conducted for the analyses shown in Table 2.

Our findings of an overall lack of detailed knowledge about the Tuskegee Syphilis Study within both Blacks and Whites (as evidence by the low scores on the TSS Facts & Myth Quiz), echo the findings of two previous reports.[18,30]

While this study refutes the notion that there is a direct connection between detailed knowledge of the Tuskegee Syphilis Study and willingness to participate in biomedical research, one limitation of this study was that it did not assess the broader question of whether and how historical events influence people's willingness to participate in research. It would be perfectly consistent with our results if the TSS had more diffuse effects on people's willingness to participate in biomedical research (e.g., a general distrust of biomedical research might be passed on after the TSS became known

without any details about the particular study having been retained). Future studies should explore this possible diffusion effect to determine if it exists, and if so, how the diffusion occurs, both within individuals and within communities.

A further limitation of this study was the post-data collection realization that the wording of one of the seven facts in the TSS Facts & Myths Quiz was not historically precise. Specifically, the fact would be more precisely worded as "The study ended once penicillin became widely available as a treatment for syphilis."

The reality that this was the first use of the TSS Facts & Myths Quiz as a measure of detailed knowledge of the TSS is yet another inherent limitation of this study. Future studies should investigate the reliability, as well as the validity, of our 7 items as a consistent and accurate measure of detailed knowledge of the TSS. Finally, space limitations prevent us from giving detailed presentation of these findings for males and for females separately, given that a previous investigation reported gender differences in awareness about the Tuskegee Syphilis Study for Whites (more males aware than females, 54% vs 38% respectively), but not for Blacks (males at 53%, female at 54%).[29]

Conclusion

The data from this follow-up analysis based upon the 1999–2000 4-City Tuskegee Legacy Project Study fail to show that being aware of the Tuskegee Syphilis Study directly affects one's willingness to participate in biomedical studies for either Blacks or for Whites, across the three cities, nearly three decades after the ending of the Tuskegee Syphilis Study. The only statistically significant finding was that in the city of Tuskegee, the historical epicenter of the index event, Blacks who had heard of the Tuskegee Syphilis Study indicated that they were willing to participate less often than Blacks who had not heard of the Tuskegee Syphilis Study. Additionally, while our previous publications revealed that Blacks were significantly more likely than Whites to have greater fear of participating (as measured by the GPFF Scale),[32,33] the data from the current report based on the same study population showed no relationship between detailed knowledge of the Tuskegee Syphilis Study and the fear of participation for either Blacks or Whites.

While there have been many discussions in the published literature, popular media, and in the community regarding the impact of the USPHS Syphilis Study at Tuskegee on the African American community,[8–21,23–34,37] it is critically important that we have a true scientific understanding of its impact on minority participation in biomedical research studies. If the Tuskegee Syphilis Study is, indeed, the reason for poor participation by minorities in biomedical studies,[23–27,30] then researchers ought to develop subject recruitment strategies for future studies that would target this issue. On the other hand, if the Tuskegee Syphilis Study it is neither the sole nor primary reason—nor after three decades even a distinctly identifiable reason—for poor levels of participation, then other strategies for are called for to ensure minority enrollment in biomedical studies.

Notes

1. Jones JH. Bad blood: the Tuskegee Syphilis Experiment. New York: Free Press, 1981, 1993.

2. Gamble VN, Fletcher JC. Chapter X. Tuskegee Syphilis Study Legacy Committee Report of 1996. In: Reverby SM, ed. Tuskegee's truths: rethinking the Tuskegee Syphilis Study. Chapel Hill and London: University of North Carolina Press, 2000.
3. Katz RV, Kegeles SS, Green BL, et al. The Tuskegee Legacy Project: history, preliminary scientific findings, and unanticipated societal benefits. Dent Clin North Am. 2003 Jan;47(1):1–19.
4. Hunninghake DB, Darby CA, Probstfield JL. Recruitment experience in clinical trials: literature summary and annotated bibliography. Control Clin Trials. 1987 Dec;8(4 Suppl):6S–30S.
5. Swanson MS, Ward AJ. Recruiting minorities into clinical trials: toward a participant-friendly system. J Natl Cancer Inst. 1995 Dec 6;87(23):1747–59.
6. Thomas SB, Quinn SC. The Tuskegee Syphilis Study, 1932 to 1972: implications for HIV education and AIDS risk education programs in the black community. Am J Public Health. 1991 Nov;81(11):1498–505.
7. Kressin NR, Meterko M, Wilson NJ. Racial disparities in participation in biomedical research. J Natl Med Assoc. 2000 Feb;92(2):62–9.
8. Shavers-Hornaday VL, Lynch CF, Burmeister LF, et al. Why are African Americans under-represented in medical research studies? Impediments to participation. Ethn Health. 1997 Mar–Jun;2(1–2):31–45.
9. Gamble VN. A legacy of distrust: African Americans and medical research. Am J Prev Med. 1993 Nov–Dec;9(6 Suppl):35–8.
10. Benedek TG. The 'Tuskegee Study' of syphilis: analysis of moral versus methodologic aspects. J Chronic Dis. 1978 Jan;31(1):35–50.
11. Corbie-Smith GM. Minority recruitment and participation in health research. N C Med J. 2004 Nov–Dec;65(6):385–7.
12. Gamble VN. The Tuskegee Syphilis Study and women's health. J Am Med Womens Assoc. 1997 Fall;52(4):195–6.
13. Pressel DM. Nuremberg and Tuaskegee: lessons for contemporary American medicine. J Natl Med Assoc. 2003 Dec;95(12):1216–25.
14. Fairchild AL, Bayer R. Uses and abuses of Tuskegee. Science. 1999 May 7;284(5416): 919–21.
15. Washington HA. Medical apartheid: the dark history of medical experimentation on Black American from colonial times to the present. New York: Doubleday Press, 2006.
16. Caplan AC. Twenty years after The Legacy of the Tuskegee Syphilis Study: when evil intrudes: the Hastings Center Repeports. Garrison: Hastings Center, 1992.
17. White RM. Misinformation and misbeliefs in the Tuskegee Study of Untreated Syphilis fuel mistrust in the healthcare system. Natl Med Assoc. 2005 Nov;97(11):1566–73.
18. Matthews AK, Sellergren SA, Manfredi C, et al. Factors influencing medical information seeking among African American Cancer patients. J. Healthcommun. 2002 May–June;7(3):205–19.
19. Bouleware LE, Cooper LA, Ratner LE, et al. Race and trust in the Elath care system. Public Health Rep. 2003 Jul–Aug;118(4):358–65.
20. Rathore SS, Krumholtz HM. Race, ethnic group, and clinical research. BMJ. 2003 Oct 4;327(7418):763–4.
21. Brandon DT, Issac LA, LaVeist TA. The legacy of Tuskegee and trust in medical care: is Tuskegee responsible for race differences in mistrust of medical care? J Natl Med Assoc. 2005 Jul;97(7):951–6.

22. U.S. Department of Health and Human Services, National Institutes of Health. National Institutes of Heath Guidelines on Inclusion of Women and Minorities as Subjects in Clinical Research. Bethesda, MD: National Institues of Health/Office of Extramural Research, 1994 Mar 18;23:(11). Available at http://grants.nih.gov/grants/funding/women_min/guidelines_update.htm.
23. Corbie-Smith GM, Thomas SB, Williams MV, et al. Attitudes and beliefs of African Americans toward participation in medical research. J Gen Intern Med. 1999 Sep; 14(9):537–46.
24. Green BL, Partridge EE, Fouad MN, et al. African-Americans attitudes regarding cancer clinical trials and research studies: results from focus group methodology. Ethn Disease. 2000;10:(1)76–86.
25. Freimuth VS, Quinn SC, Thomas SB, et al. African Americans' views on research and the Tuskegee Syphilis Study. Soc Sci Med. 2001 Mar;52(5):797–808.
26. Bates BR, Harris TM. The Tuskegee Study of Untreated Syphilis and public perceptions of biomedical research: a focus group study. J Natl Med Assoc. 2004 Aug; 96(8):1051–64.
27. Sengupta S, Strauss RP, DeVellis R, et al. Factors affecting African-American participation in AIDS research. J Acquir Immune Defic Syndr. 2000 Jul 1;24(3):275–84.
28. Brown DR, Topcu M. Willingness to participate in clinical treatment research among older African Americans and Whites. Gerontologist. 2003 Feb;43(1):62–72.
29. Green BL, Maisiak R, Wang MQ, et al. The Tuskegee Legacy Project: willingness of minorities to participate in biomedical research. J Health Care Poor Underserved. 2006 Nov;17(4):698–715.
30. Shavers VL, Lynch CF, Burmeister LF. Knowledge of the Tuskegee study and its impact on the willingness to participate in medical research studies. J Natl Med Assoc. 2000 Dec;92(12):563–72.
31. McCallum JM, Arekere DM, Green BL, et al. Awareness and knowledge of the U.S. Public Health Service syphilis study at Tuskegee: implications for biomedical research. J Health Care Poor Underserved. 2006 Nov;17(4):716–33.
32. Katz RV, Kegeles SS, Kressin NR, et al. The Tuskegee Legacy Project: willingness of minorities to participate in biomedical research. J Health Care Poor Underserved. 2006 Nov;17(4):698–715.
33. Katz RV, Green BL, Kressin NR, et al. Willingness of minorities to participate in biomedical studies: confirmatory findings from a follow-up study using the Tuskegee Legacy Project Questionnaire. J Natl Med Assoc. 2007 Sep;99(9):1052–60. Erratum in: J Natl Med Assoc. 2007 Nov;99(11):1216.
34. Katz RV, Kegeles SS, Kressin NR, et al. Awareness of the Tuskegee Syphillis Study and the U.S. Presidential Apology and their influence on minority participation in biomedical research. Am J Public Health. 2008 June;98:1137–42.
35. SPSS Inc. SPSS for Windows 14.0. Chicago, IL: SPSS Inc., 2006. Available at http://www.spss.com/.
36. SAS Institue Inc. SAS 9.1. Help and documentation. Cary, NC: SAS Institute Inc., 200–2004.
37. Echtenkamp-Klein J, Fletcher JC (organizers). Doing bad in the name of good: the legacy of the Tuskegee Syphilis Study. Presented at: University of Virginia Bioethics Conference, Charlottesville (VA), Apr 1994. Available at http://www.hsl.virginia.edu/historical/medical_history/bad_blood/symposium.cfm.

Awareness of the Tuskegee Syphilis Study and the US Presidential Apology and Their Influence on Minority Participation in Biomedical Research

Ralph V. Katz, DMD, PhD, MPH, S. Stephen Kegeles, PhD, Nancy R. Kressin, PhD, B. Lee Green, PhD, Sherman A. James, PhD, Min Qi Wang, PhD, Stefanie L. Russell, DDS, PhD, MPH, and Cristina Claudio, PhD

Objectives. We compared the influence of awareness of the Tuskegee Syphilis Study and the presidential apology for that study on the willingness of Blacks, non-Hispanic Whites, and Hispanics to participate in biomedical research.

Methods. The Tuskegee Legacy Project Questionnaire was administered to 1133 adults in 4 US cities. This 60-item questionnaire addressed issues related to the recruitment of minorities into biomedical studies.

Results. Adjusted multivariate analysis showed that, compared with Whites, Blacks were nearly 4 times as likely to have heard of the Tuskegee Syphilis Study, more than twice as likely to have correctly named Clinton as the president who made the apology, and 2 to 3 times more likely to have been willing to participate in biomedical studies despite having heard about the Tuskegee Syphilis Study (odds ratio [OR] = 2.9; 95% confidence interval [CI] = 1.4, 6.2) or the presidential apology (OR = 2.3; 95% CI = 1.4, 3.9).

Conclusions. These marked differences likely reflect the cultural reality in the Black community, which has been accustomed to increased risks in many activities. For Whites, this type of information may have been more shocking and at odds with their expectations and, thus, led to a stronger negative impact. (*Am J Public Health.* 2008;98:1137–1142. doi:10.2105/AJPH.2006.100131)

The US Public Health Service Tuskegee Syphilis Study (1932–1972) is arguably the most infamous biomedical research study in US history.[1–5] This study enrolled 399 Black sharecroppers in Macon County, Ala, and studied the effects of not treating their syphilis.[6,7] There is widespread belief that the "legacy" of this unethical study is that the Black community has a greater reluctance to participate in clinical research studies because of the abuses foisted on the participants in that study. Although a considerable amount has been written about the long-lasting effects of the Tuskegee Syphilis Study on the Black community, most of this work has been from a legal, historical, ethical, or access to health care perspective.[8–20]

Surprisingly few research articles have directly examined whether any differential participation of Blacks or other minorities in biomedical studies compared with participation of Whites was because of the legacy of the Tuskegee Syphilis Study or because of other factors.[21–28] A recent literature review noted that only 5 of the published studies to date have presented quantified data that compared Black with White participation and the relation to the Tuskegee Syphilis Study,[29] and most of them only used a single question on willingness to participate as their measure of this complex decison.[30] A recent study used a series of questions to create 2 validated scales to measure willingness to participate and found that Blacks self-reported that despite having a higher fear of participation, they were just as likely as Whites to participate in biomedical research.[30]

For our study, we used data from a telephone survey of adults in 4 US cities that used the Tuskegee Legacy Project (TLP) Questionnaire.[30] We sought to compare racial/ethnic differences among Blacks, Whites, and Hispanics, in the level of awareness of the Tuskegee Syphilis Study and the US presidential apology made to the Black community, and to compare the self-reported influence among Blacks, Whites, and Hispanics of both the study and the presidential apology on the willingness to participate in biomedical studies.

METHODS

The TLP Questionnaire was administered via random-digit-dialed telephone interviews to 1133 Blacks, Hispanics, and Whites aged 18 years and older in 4 city and county areas: Birmingham and Jefferson County, and Tuskegee and Macon County, Ala; Hartford and Hartford County, Conn; and San Antonio and Bexar County, Tex. All interviews were conducted between March 1999 and November 2000.

The TLP Questionnaire, a 60-item instrument, addresses a range of issues related to the recruitment of minorities into biomedical studies. Details on the history and development of the TLP Questionnaire and justifications of the methodological decisions both for the selection of the 4 cities and for the analysis of the TLP Questionnaire have been described elsewhere.[3,30] All TLP Questionnaire interviews were conducted in English. Respondents answered questions but were provided no information about the Tuskegee Syphilis Study during the interview.

The interviews were administered by the Survey Research Unit of the University of Alabama at Birmingham. The target population consisted of noninstitutionalized adults who lived in households with working telephones in the 4 targeted cities and counties. The sample of households in each of the 4 locations was supplied by Survey Sampling Inc (Fairfield, Conn) and was based upon a simple random sampling of telephone numbers that used the 3-digit telephone exchanges for those local calling areas with partial screening

for nonworking or business numbers. Thirteen interviewers, trained for the survey, used the full computer-assisted telephone interviewing technology. Unresolved numbers were retired after 20 attempts. Interviewers were supervised at all times and randomly electronically monitored a minimum of 4 times per month.

We conducted unadjusted bivariate analyses, which were followed by multivariate logistic regression analyses, adjusted for age, gender, education, and income, as well as city. To acknowledge and account for cultural differences among the cities (i.e., above and beyond simple demographic differences), we included the variable *city* as a separate covariate in all multivariate analyses of the study sample as a whole. We conducted the bivariate and multivariate analyses using SPSS version 14.0 (SPSS Inc, Chicago, Ill) and SAS version 9.0 (SAS Institute Inc, Cary, NC). We calculated confidence intervals (CIs) for percentages using Stata version 9 (Stata Corp, College Station, Tex) with its module Confidence Interval for Proportions.

RESULTS

Response rates in Birmingham and Jefferson County, Tuskegee and Macon County, Hartford and Hartford County, and San Antonio and Bexar County were 70%, 65%, 49%, and 50%, respectively. The overall completion rate (number of completed interviews per number of initiated interviews) exceeded 90% in each city. For San Antonio, the major Spanish-speaking Hispanic population accessed in this survey, 10% of the contacted individuals indicated that they could not participate with the English-language-only instrument. Table 1 shows the age, gender, education, and income distribution of the 1133 respondents within the 3 racial/ethnic groups.

To determine if a respondent was aware of the Tuskegee Syphilis Study, the TLP Questionnaire had 2 separate recognition probes. The first recognition probe consisted of the respondents being asked directly whether they had ever heard about the Tuskegee Syphilis Study. The 3 racial/ethnic groups differed markedly on responses to this first recognition probe for the Tuskegee Syphilis Study with 72.6% of Blacks, 55.2% of Whites, and 23.6% of Hispanics answering yes ($P<.001$, by the χ^2 test). Each contrast between any 2 of the racial/ethnic groups was also statistically significant at $P<.001$. The second probe, which occurred 6 questions later in the interview, was only asked of individuals who either said no to having heard of the Tuskegee Syphilis Study in the first probe or who answered yes to that first probe but could provide no details about that study.

Of the original 1133 respondents who were interviewed, 57.2% (95% CI=54.3%, 60.1%) indicated that they had heard of the Tuskegee Syphilis Study when the data from both probes were combined. For the 1128 respondents who filtered through the 2 probes with valid responses (i.e., a yes or no response), there were marked differences among the racial/ethnic groups on their final yes or no answer to the inquiry on whether they had ever heard of the Tuskegee Syphilis Study.

Results of an unadjusted bivariate analysis based upon the 2-probe combination showed that 76.4% of the 353 Blacks in the study (95% CI=71.7%, 80.8%), 56.8% of the 623 Whites in the study (95% CI=52.8%, 60.8%), and 25.3% of the 157 Hispanics in the study (95% CI=18.9%, 33.0%) had indicated that they had heard of the Tuskegee Syphilis Study (Figure 1). This difference was statistically significant across the 3 racial/ethnic groups ($P<.001$, by the χ^2 test), with each 2-way contrast also statistically significant ($P<.001$, by the χ^2 test). Across racial/ethnic groups, a comparison of yes responses to the first probe only versus the yes responses to the 2-probe combination revealed that only a very slight upward correction factor resulted from the use of the second probe for each of the 3 racial/ethnic groups (i.e., an increase of 3.8 percentage points in Blacks, 1.6 percentage points in Whites, and 1.7 percentage points in Hispanics).

Figure 1 also shows the percentage of correct responses to: "Has any US President ever apologized for the Tuskegee Syphilis Study?" and "Which US president?" Of the Blacks in the study, 42.5% had heard of the presidential apology (95% CI=27.3%, 47.8%) and 34.8% knew that President Clinton had given the apology (95% CI=29.9%, 40.0%). Of the Whites in the study, 28.1% had heard of the presidential apology (95% CI=24.6%, 31.8%) and 24.1% knew that President Clinton had given the apology (95% CI=20.8%, 27.6%). Of the Hispanics in the study, 8.3% had heard of the presidential apology (95% CI=4.5%, 13.7%) and 3.2% knew that President Clinton had given the apology (95% CI=1.0%, 7.3%).

Table 2 shows the multivariate logistic regression analyses of the responses to the

TABLE 1—Age, Gender, Education, and Income Distribution of Respondents (N = 1113), by Racial/Ethnic Group: Tuskegee Legacy Project Study, 1999–2000

	Blacks[a,b]	Whites[a,c]	Hispanics[b,c]
Total, No.	353	623	157
Age, mean (SD)	49.1 (16.5)	53.8 (17.0)	41.5 (16.1)
Male, % (95% CI)	52.1 (46.8, 57.4)	48.3 (44.3, 52.3)	39.5 (31.8, 47.6)
Education level, % (95% CI)			
Less than high school graduate	21.6 (17.4, 26.4)	11.8 (9.4, 14.7)	14.0 (9.0, 20.4)
High school graduate or some college	60.5 (55.3, 65.8)	51.3 (47.4, 55.4)	61.0 (53.5, 68.8)
College graduate or higher	17.9 (14.0, 22.2)	36.9 (33.1, 40.8)	25.0 (16.3, 32.4)
Income level, % (95% CI)			
< $20 000	42.8 (37.6, 48.1)	21.3 (18.2, 24.8)	41.7 (33.6, 49.5)
$20 000–$74 999	52.1 (46.8, 57.4)	58.4 (54.4, 62.3)	52.5 (44.1, 60.3)
≥ $75 000	5.1 (3.1, 7.9)	20.3 (17.1, 23.6)	5.8 (2.7, 10.6)

Note. CI = confidence interval.
[a]Statistically significant contrasts for Blacks versus Whites: differed on age, education, and income ($P \leq .05$).
[b]Statistically significant contrasts for Blacks versus Hispanics: differed on age and gender ($P \leq .05$).
[c]Statistically significant contrasts for Hispanics versus Whites: differed on age, gender, education, and income ($P \leq .05$).

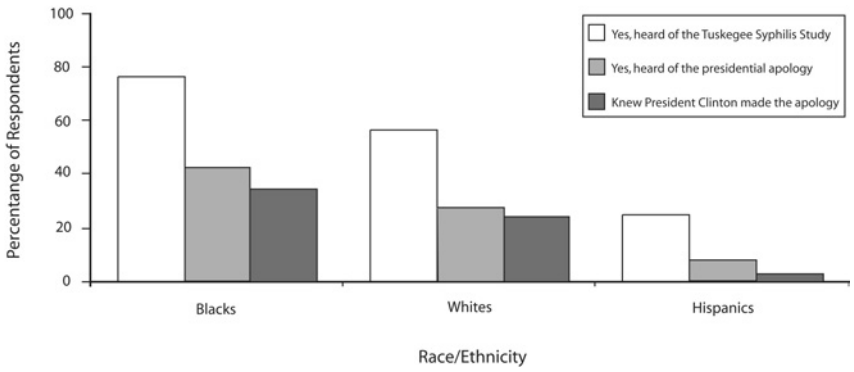


FIGURE 1—Responses to having heard of the Tuskegee Syphilis Study and the presidential apology, and knowing who made that apology, among Blacks, Whites, and Hispanics: Tuskegee Legacy Project Study, 1999–2000.

same 3 questions adjusted for age, gender, education, income, and city. The adjusted multivariate analysis shows that the odds of hearing of the Tuskegee Syphilis Study were nearly 4 times greater for Blacks than for Whites (odds ratio [OR]=3.9; 95% CI=2.6, 5.7). The difference in odds between Hispanics and Whites was not statistically significant (OR=1.6; 95% CI=0.93, 2.7). The adjusted multivariate analysis for having heard of the presidential apology showed that whereas the odds for Blacks did not significantly differ from the odds for Whites (OR=1.6; 95% CI=0.91, 2.7), the odds of correctly naming Clinton as the President who made the apology were 2 times greater for Blacks than for Whites (OR=2.3; 95% CI=1.6, 3.4). Too few Hispanics had heard of the apology to include Hispanics in these latter 2 analyses.

Given that the ORs for these 3 questions by city (Table 2) showed a strong effect of city of residence on awareness of the study and of the presidential apology, adjusted ORs were computed to assess the racial difference for each of the 3 questions within each city. Because of the limitations of sample size on the stability and interpretation of data, this further analysis was only conducted for Blacks and Whites in the 3 cities with substantial numbers of these 2 racial groups (i.e., Birmingham, Tuskegee, and Hartford). The results of this additional within-city analysis indicated that the odds of Blacks having ever heard of the Tuskegee Syphilis Study were 4 to 5 times higher than were the odds for Whites in the cities of Hartford and Birmingham (OR=4.37 [95% CI=2.40, 7.96] and OR=5.47 [95% CI=2.35, 12.71], respectively), but not significantly different from Whites in the city of Tuskegee (OR=2.37; 95% CI=0.89, 6.34).

Although there were no statistically significant differences in odds between Blacks and Whites for the question "Has any US President ever apologized for the Tuskegee Syphilis Study?" in any of the 3 cities, statistically significant differences were observed for the third question: "Which US President?" For each of the 3 cities, Blacks were 2.5 to 3 times more likely than were Whites to correctly name Clinton as the president who made the apology (Hartford: OR=2.81 [95% CI=1.31, 6.03]; Birmingham: OR=2.97 [95% CI=1.44, 6.11]; and Tuskegee: OR=2.50 [95% CI=1.33, 4.70]).

Respondents who replied yes to the 2-probe series on ever having heard of the Tuskegee Syphilis Study were then asked the follow-up question: "As a result of what you have heard about the Tuskegee Syphilis Study, how likely are you to participate in a medical research study?" with responses on a 5-point Likert scale ranging from *much more likely* to *much less likely.* Because very few Hispanic respondents had responded yes to this 2-probe series (n=39), this analysis was performed only for Black and White respondents.

Figure 2 shows an unadjusted bivariate analysis that revealed that among those who had heard of the Tuskegee Syphilis Study, the negative influence on the likelihood of participation in future studies was less among Blacks than among Whites; i.e., 50.0% of Blacks and 70.3% of Whites reported that they were less likely to participate as a result of what they had heard about the Tuskegee Syphilis Study ($P<.001$; Kendall's tau-B).

Multivariate logistic regression analysis of this question (adjusted for age, gender, education, income, and city) revealed that despite what they had heard about the Tuskegee Syphilis Study, the odds of Blacks indicating a willingness to participate in biomedical studies were nearly 3 times greater than the odds of Whites (OR=2.9; 95% CI=1.4, 6.2).

After the respondent was asked "Has any US President ever apologized for the Tuskegee Syphilis Study?" a follow-up question was asked: "Based upon what you heard about the apology, would it influence your decision to join a biomedical research study today? Did that apology make you more or less likely to join a study?" This question also had a 5-point Likert scale of responses ranging from *much more likely* to *much less likely.*

The results from an unadjusted bivariate analysis of this follow-up question are shown in Figure 2 and again revealed a much less negative influence on Blacks, with 41.3% (vs 61.8% of Whites) self-reporting that they were less likely to join a biomedical study as a result of the presidential apology ($P=.008$, by the χ^2 test). Multivariate logistic regression analysis of this question (adjusted for age, gender, education, income, and city) revealed that the odds of indicating that they were more likely to participate in biomedical studies as a result of having heard of the presidential apology were more than 2 times greater for Blacks than for Whites (OR=2.3; 95% CI=1.4, 3.9).

DISCUSSION

Aspects of Current Study

This study clearly shows marked and statistically significant differences among Blacks, Whites, and Hispanics with regard to their awareness of the Tuskegee Syphilis Study and the presidential apology for that study. The 3 key questions analyzed in this report (Have you ever heard of the Tuskegee Syphilis

TABLE 2—Logistic Regression Multivariate Analyses for Key Questions on Questionnaire: Tuskegee Legacy Project Study, 1999–2000

	"Have You Ever Heard of the Tuskegee Syphilis Study?" OR (95% CI)	"Has Any US President Ever Apologized for the Tuskegee Syphilis Study?" OR (95% CI)	"Which US President?" OR (95% CI)
Race/ethnicity[a]			
Blacks	3.87 (2.63, 5.70)	1.58 (0.91, 2.72)	2.33 (1.63, 3.80)
Hispanics	1.59 (0.93, 2.73)	. . .[b]	. . .[b]
Education levels[c]			
High school graduate	1.36 (0.84, 2.20)	1.30 (0.58, 2.94)	1.36 (0.79, 2.34)
Some college	2.20 (1.28, 3.78)	1.70 (0.71, 4.11)	2.05 (1.14, 3.68)
College graduate	3.14 (1.73, 5.69)	2.19 (0.83, 5.79)	3.62 (1.91, 6.85)
Higher than college graduate	5.89 (2.90, 11.99)	2.16 (0.72, 6.52)	3.49 (1.70, 7.17)
Income levels,[d] $			
20 000–34 999	1.29 (0.85, 1.96)	0.74 (0.37, 1.47)	1.60 (1.00, 2.55)
35 000–49 999	1.53 (0.93, 2.52)	1.18 (0.51, 2.71)	2.60 (1.50, 4.52)
50 000–74 999	2.20 (1.28, 3.77)	1.66 (0.66, 4.14)	2.97 (1.65, 5.36)
≥ 75 000	1.66 (0.94, 2.92)	1.10 (0.42, 2.88)	1.89 (0.97, 3.58)
Gender[e]			
Female	0.66 (0.48, 0.90)	0.99 (0.00, 1.63)	0.61 (0.43, 0.87)
City[f]			
Birmingham, Ala	3.34 (2.17, 5.12)	1.93 (0.97, 3.84)	3.32 (2.07, 5.32)
Tuskegee, Ala	16.76 (10.02, 28.02)	3.24 (1.66, 6.30)	7.81 (4.97, 12.29)
San Antonio, Tex	0.71 (0.45, 1.10)	0.62 (0.22, 1.74)	0.44 (0.20, 0.99)
Age, y	1.02 (1.01, 1.03)	1.01 (1.00, 1.03)	1.02 (1.01, 1.03)

Notes. OR = odds ratio; CI = confidence intervals. Analyses of a "yes" answer to the first 2 questions and a correct answer to the third, open-ended question. Analyses adjusted for race, age, gender, education, income, and city.
[a]Reference group was Whites.
[b]Hispanics excluded from analysis because of too few respondents to this question.
[c]Reference group was less than high school graduate.
[d]Reference group was less than $20 000.
[e]Reference group was male.
[f]Reference group was Hartford, Conn.

Study? Has any US President ever apologized for the Tuskegee Syphilis Study? Which US President made the apology?) led to the conclusions that Blacks were much more aware of the Tuskegee Syphilis Study than either Whites or Hispanics, and that the odds of knowing that President Clinton had made the apology were much higher for Blacks than for Whites. But Blacks and Whites did not differ significantly on knowing that a presidential apology had been made.

The responses to these 3 questions showed a much stronger impact of the Tuskegee study in the 2 southern cities of Birmingham and Tuskegee than in the northern city of Hartford, as shown in Table 2. The odds ranged up to nearly 17-times greater for Tuskegee than for Hartford. Interestingly, although the intracity adjusted analyses for each of these 3 cities by race revealed that Blacks were more likely than were Whites to know that President Clinton had made the apology for that study, a racial difference was only observed in the 2 cities of Birmingham and Hartford (not in Tuskegee) for the question "Have you ever heard of the Tuskegee Syphilis Study?" Because the city of Tuskegee is the historical epicenter of this issue and its legacy, this seemingly unexpected result is likely because of an exceptionally high rate of having heard of the Tuskegee Syphilis Study among both Whites and Blacks in the city of Tuskegee (i.e., about 90% for both Blacks and Whites in Tuskegee).

Overall, because nearly 60% of the Black respondents had heard of the Tuskegee Syphilis Study (ranging from 89.8% to 71.1% to 46.5% to 29.6% in Tuskegee, Birmingham, Hartford, and San Antonio, respectively) health disparities researchers working in the Black communities in the future must acknowledge this level of awareness. Specifically, for researchers to create a respectful, comfortable, and inviting atmosphere for all potential participants when planning and recruiting participants into studies, the researchers must take this overall high level of awareness into account.

Because Blacks overall were much more aware of the Tuskegee Syphilis Study and generally knew more details about the apology than did Whites and Hispanics, the impact of this increased awareness of the study in Blacks is most interesting. Overall, the plurality of both Blacks and Whites who had heard of the Tuskegee Syphilis Study indicated that they were less likely to participate in biomedical studies, but this negative impact was far more pronounced among Whites (71%) than among Blacks (50%). Hence, Blacks who had heard of the Tuskegee Syphilis Study were much less negatively affected by that awareness than were Whites. This was true regardless of whether it was the impact of having heard about the Tuskegee Syphilis Study or of having heard of the presidential apology for the study.

Conversely, Whites' willingness to participate in biomedical research studies was much more negatively affected by awareness of each of these events. This marked difference in the observed impact may reflect the daily cultural reality in the Black community, which has for a long time been accustomed to increased risks for Blacks in many activities. For Whites, this type of information (or news) may have been more shocking and at odds with their daily expectations and, thus, led to a stronger negative impact on their future decisionmaking.

The major methodological finding of this study, which resulted from the use of 2 recognition probes to determine whether a respondent was aware (i.e., recognized the name) of the Tuskegee Syphilis Study, provides clear evidence that a single probe suffices when one is asking about awareness of the

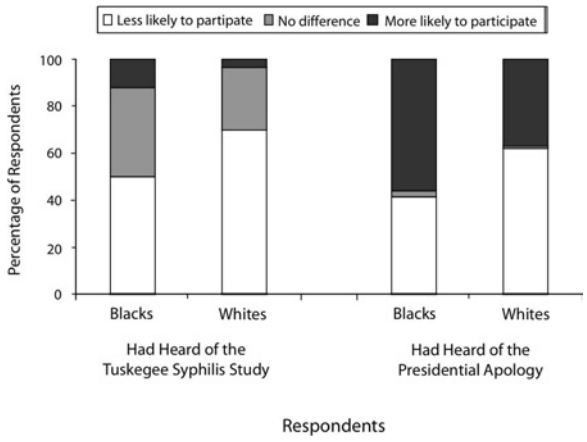


FIGURE 2—Percentage of respondents having heard about the Tuskegee Syphilis Study and about the presidential apology for the study and the influence of both on the likelihood to participate in biomedical studies, by race/ethnicity: Tuskegee Legacy Project Study, 1999–2000.

Tuskegee Syphilis Study. The second probe did slightly increase the number of respondents in each of the racial/ethnic groups that indicated awareness of the study—by only 3.8, 1.7, and 1.6 percentage points for Blacks, Hispanics, and Whites, respectively. These findings suggest that use of a single recognition probe, as was done in all the prior studies that investigated the Tuskegee Syphilis Study, is quite accurate and would be sufficient for future studies.

Comparison With Prior Literature

Five published studies have reported both on having heard of the Tuskegee Syphilis Study and on willingness to participate in biomedical research by race.[18,25–28] Across these 5 studies, the percentage of Blacks who had heard of the Tuskegee Syphilis Study ranged from 42% to 81%; for Whites the range was between 18% and 46%. Our findings are at the high end of the percentages reported to date, both for Blacks (73% on our first probe, 76% after the second probe) and for Whites (55% on our first probe, 57% after the second probe). No ready explanation for this observation is provided by demographic differences (such as age, gender, education, or income) between our study population and those previous studies.

Although the professional literature related to health care in the United States is replete with articles that refer to the impact and assumed legacy of the Tuskegee Syphilis Study,[1–3,6,13–24,28–30] perhaps the most unusual measure of the depth of cultural influence of this legacy is documented by its recent appearance as the core theme of a Marvel comic book 7-issue series: "Truth: Red, White and Black," written in 2003 as a presequel to the Captain America series. In this fictional presequel series, research abuses abound in experiments done on a Black military unit. An injected compound is used to biomedically develop a "supersoldier," one that once perfected on this "research abused" Black military unit is then used safely to create Captain America, a White supersoldier of comic book fame.[31]

Conclusions

Our study of the reputed legacy of the Tuskegee Syphilis Study reports on the largest and most geographically diverse study sample to date. This study is also the first to quantitatively report on the community impact—among Blacks, Whites, and Hispanics—of President Clinton's 1997 apology for the Tuskegee Syphilis Study. This apology was made to the Black community at large, as well as directly to surviving study participants and the families of the nonsurvivors.

Our findings provide clear evidence that Blacks were both much more likely than Whites or Hispanics to have heard of the Tuskegee Syphilis Study and to know that President Clinton had made the apology. Most interestingly, despite Blacks being more aware of both the Tuskegee Syphilis Study and who made the presidential apology, Whites who had heard either of the study or of the presidential apology were more negatively influenced toward participation in biomedical research than were Blacks who had heard of either event.

These findings, plus the regional differences observed between the northern city of Hartford and the southern cities of Birmingham and Tuskegee, strongly suggest that if future studies are to attain their goal of having a diverse set of study participants as mandated by federal research guidelines, investigators who conduct clinical and community-based studies in the future need to recognize and incorporate these racial/ethnic, geographical, and cultural differences into their recruitment and retention plans. ■

About the Authors

Ralph V. Katz is with the Department of Epidemiology and Health Promotion, New York University College of Dentistry, New York. At the time of the study, S. Stephen Kegeles was with the Department of Behavioral Sciences and Community Health, School of Dental Medicine, University of Connecticut, Farmington. Nancy R. Kressin is with the Department of General Internal Medicine, Boston University Medical School, Boston, Mass, and the Center for Health Quality, Outcomes and Economic Research, Department of Veterans Affairs, Bedford, Mass. B. Lee Green is with the Office of Institutional Diversity, H. Lee Moffitt Cancer Center and Research Institute, Tampa, Fla. Sherman A. James is with the Terry Sanford Institute of Public Policy, Duke University, Durham, NC, and the Department of Community and Family Medicine, Duke University, Durham. Min Qi Wang is with the Department of Public and Community Health, School of Public Health, University of Maryland, College Park. Stefanie L. Russell is with the Department of Epidemiology and Health Promotion, College of Dentistry, New York University, New York. Cristina Claudio is with the University of Puerto Rico School of Dentistry, Medical Sciences Campus, San Juan.

Requests for reprints should be sent to Ralph V. Katz, DMD, MPH, PhD, Professor and Chair, Department of Epidemiology & Health Promotion, NYU College of Dentistry, 345 E 24th St, MC-9416, New York, NY 10010 (e-mail: ralph.katz@nyu.edu).

This article was accepted February 5, 2007.

Contributors

R.V. Katz originated and directed the study, led the development of the Tuskegee Legacy Project Questionnaire and the data analysis, wrote the initial draft of the article, and led the team's crafting of the final article. S.S. Kegeles, N.R. Kressin, S.A. James, and B.L. Green

developed the Tuskegee Legacy Project Questionnaire, assisted with the writing of the grants that supported this research, helped plan the data analysis and data interpretation, and contributed to the writing of the final article. M.Q. Wang and S.L. Russell conducted the statistical analyses, helped to plan and finalize the data interpretation and contributed to the writing of the final article. C. Claudio helped plan the data analysis and data interpretation, and contributed to the writing of the final article.

Acknowledgments

This study was supported by the National Institutes of Dental and Craniofacial Research at the National Institutes of Health: the Northeastern Minority Oral Health Research Center, a Regional Research Center for Oral Health (P50 DE10592), and the New York University Oral Cancer Research on Adolescent and Adult Health Promotion Center, an Oral Health Disparities Center (U54 DE014257).

We wish to acknowledge the invaluable and senior leadership of our deceased colleague, Stephen Kegeles, in the design of both the Tuskegee Legacy Project Questionnaire and Tuskegee Legacy Project Study, as well in the data collection phase, of this study. Because of his death, we dearly missed his insights during the data interpretation phase.

Human Participant Protection

This study was approved by the institutional review boards of the University of Connecticut Health Center and of New York University.

References

1. Jones JH. *Bad Blood: The Tuskegee Syphilis Experiment.* New York, NY: Free Press; 1981.

2. Tuskegee Syphilis Study Legacy Committee Report of 1996, Centers for Disease Control and Prevention Workshop at Tuskegee University. In: Reverby SM, ed. *Tuskegee's Truths: Rethinking the Tuskegee Syphilis Study.* Chapel Hill: University of North Carolina Press; 2000.

3. Katz RV, Kegeles SS, Green LB, Kressin NR, James SA, Claudio C. The Tuskegee Legacy Project: history, preliminary scientific findings, and unanticipated societal benefits. *Dent Clin North Am.* 2003;47: 1–19.

4. Hunninghake DB, Darby CA, Probstfield JL. Recruitment experience in clinical trials: literature summary and annotated bibliography. *Control Clin Trials.* 1987; 8(suppl 4)6S–30S.

5. Swanson MS, Ward AJ. Recruiting minorities into clinical trials: toward a participant-friendly system. *J Natl Cancer Inst.*1995;87:1747–1759.

6. Thomas SB, Quinn SC. The Tuskegee Syphilis Study, 1932 to 1972: implications for HIV education and AIDS risk education programs in the Black community. *Am J Public Health.* 1991;81:1498–1505.

7. Kressin NR, Meterko M, Wilson NJ. Racial disparities in participation in biomedical research. *J Natl Med Assoc.* 2000;92:62–69.

8. Shavers-Hornaday VL, Lynch CF, Burmeister LF, Torner JC. Why are African Americans under-represented in medical research studies? Impediments to participation. *Ethn Health.* 1997;2:31–45.

9. Gamble VN. A legacy of distrust: African Americans and medical research. *Am J Prev Med.* 1993; 9:(suppl 6)35–38.

10. Caplan AL. Twenty years after. The legacy of the Tuskegee Syphilis Study. When evil intrudes. *Hastings Cent Rep.* 1992;22:29–32.

11. Benedek TG. The "Tuskegee Study" of syphilis: analysis of moral versus methodologic aspects. *J Chron Dis.* 1978;31:35–50.

12. Corbie-Smith GM. Minority recruitment and participation in health research. *N C Med J.* 2004;6: 385–387.

13. Gamble VN. The Tuskegee Syphilis Study and women's health. *J Am Med Womens Assoc.* 1997;52: 195–196.

14. Matthews AK, Sellergren SA, Manfredi C, Williams M. Factors influencing medical information seeking among African American cancer patients. *J Health Commun.* 2002;7:205–219.

15. Pressel DM. Nuremberg and Tuskegee: lessons for contemporary American medicine. *J Natl Med Assoc.* 2003;95:1216–1225.

16. Rathore SS, Krumholz HM. Race, ethnic group, and clinical research. *BMJ.* 2003;327:763–764.

17. Boulware LE, Cooper LA, Ratner LE, LaVeist TA, Powe NR. Race and trust in the health care system. *Public Health Rep.* 2003;118:358–365.

18. Brandon DT, Isaac LA, LaVeist TA. The legacy of Tuskegee and trust in medical care: is Tuskegee responsible for race differences in mistrust of medical care*? J Natl Med Assoc.* 2005;97:951–956.

19. White RM. Misinformation and misbeliefs in the Tuskegee Study of Untreated Syphilis fuel mistrust in the healthcare system. *J Natl Med Assoc.* 2005;97: 1566–1573.

20. Fairchild AL, Bayer R. Uses and abuses of Tuskegee. *Science.* 1999;284:919–921.

21. Corbie-Smith G, Thomas SB, Williams MV, Moody-Ayers S. Attitudes and beliefs of African Americans toward participation in medical research. *J Gen Intern Med.* 1999;14:537–546.

22. Green BL, Partridge EE, Fouad MN, Kohler C, Crayton EF, Alexander L. African-American attitudes regarding cancer clinical trials and research studies: results from focus group methodology. *Ethn Dis.* 2000; 10:76–86.

23. Freimuth VS, Quinn SC, Thomas SB, Cole G, Zook E, Duncan T. African Americans' views on research and the Tuskegee Syphilis Study. *Soc Sci Med.* 2001;52:797–808.

24. Bates BR, Harris TM. The Tuskegee Study of Untreated Syphilis and public perceptions of biomedical research: a focus group study. *J Natl Med Assoc.* 2004; 96:1051–1064.

25. Sengupta S, Strauss RP, DeVellis R, Quinn SC, DeVellis B, Ware WB. Factors affecting African-American participation in AIDS research. *J Acquir Immune Defic Syndr.* 2000;24:275–284.

26. Brown DR, Topcu M. Willingness to participate in clinical treatment research among older African Americans and Whites. *Gerontologist.* 2003;43:62–72.

27. Green BL, Maisiak R, Wang MQ, Brill MF, Ebeling N. Participation in health education, health promotion, and health research by African Americans: effects of the Tuskegee Syphilis Experiment. *J Health Educ.* 1997;28:196–201.

28. Shavers VL, Lynch CF, Burmeister LF. Knowledge of the Tuskegee study and its impact on willingness to participate in medical research studies. *J Natl Med Assoc.* 2000;92:563–572.

29. McCallum JM, Arekere DM, Green BL, Katz RV, Rivers BM. Awareness and knowledge of the U.S. Public Health Service syphilis study at Tuskegee: implications for biomedical research. *J Health Care Poor Underserved.* 2006;17:716–733.

30. Katz RV, Kegeles SS, Kressin NR, et al. The Tuskegee Legacy Project: willingness of minorities to participate in biomedical research. *J Health Care Poor Underserved.* 2006;17:695–715.

31. Morales R, Baker K. *Truth: Red, White and Black, Vol I–VII, a Pre-sequel to the Captain America Series.* New York, NY: Marvel Comics; January–July 2003.

Participation in Biomedical Research Studies and Cancer Screenings: Perceptions of Risks to Minorities Compared With Whites

Ralph V. Katz, DMD, MPH, PhD, Min Qi Wang, PhD, B. Lee Green, PhD, Nancy R. Kressin, PhD, Cristina Claudio, PhD, Stefanie Luise Russell, DDS, MPH, PhD, and Christelle Sommervil, MPH, MS, CCRP

Cancer, Culture and Literacy

Background: *This analysis was conducted to determine whether there is a difference among blacks, Hispanics, and whites in their perception of risks associated with participating in either a biomedical study or a cancer screening.*

Methods: *The Tuskegee Legacy Project Questionnaire, which focused on research subject participation, was administered in two different surveys (1999–2000 and 2003) in seven cities. The Cancer Screening Questionnaire was administered in 2003 in three cities.*

Results: *The study sample across the three surveys consisted of 1,064 blacks, 781 Hispanics, and 1,598 non-Hispanic whites. Response rates ranged from 44% to 70% by city. Logistic regression analyses, adjusted for age, sex, education, income, and city, revealed that blacks and Hispanics each self-reported that minorities, compared with whites, are more likely to be "taken advantage of" in biomedical studies and much less likely to get a "thorough and careful examination" in a cancer screening (odds ratios ranged from 3.6 to 14.2).*

Conclusions: *Blacks and Hispanics perceive equally high levels of risk for participating in cancer screening examinations and for volunteering to become research subjects in biomedical studies. This perception provides a strong message about the need to overtly address this critical health disparities issue.*

Introduction

Two widely espoused but widely different goals within the field of health disparities are (1) to increase the number of minorities who volunteer to participate in biomedical research studies, including clinical trials, and (2) to increase the number of minorities who participate in screening programs that target the early detection of cancer.[1] The first goal will ensure that the new treatments and preventive methods work equally well for minorities, where the burden of health disparities resides, in keeping with the letter as well as the spirit of the 1994 law for the Inclusion of Women and Minorities in biomedical studies.[2] The second goal is intended to enhance the likelihood of ensuring timely treatment and longer survival for minority persons with cancer, since again minorities disproportionally bear the health disparities burden for this deadly disease.[3-6]

The well-established literature addressing trust/distrust issues in biomedical research has focused largely on blacks in the United States and has led to the conclusion that blacks, compared with whites, have much lower trust related to biomedical research based on surveys that have directly asked questions about "degree of trust" in their survey instruments.[7-18] Other surveys that have targeted their questions not directly on the broad issue of trust per se but on the closely related, narrower topics of willingness to participate in biomedical studies and/or on fear of participation in biomedical studies have generally found that blacks reported equal willingness to participate but with higher levels of fear of participation.[19-35] More recently, there is a growing body of literature that has identified similar issues of lower trust related to biomedical services within Hispanic communities in the United States.[36-40]

While there is clear evidence that minorities have lower cancer screening rates than whites,[36-38] most of

From the Department of Epidemiology & Health Promotion at the New York University College of Dentistry, New York, New York (RVK, SLR); the Department of Public & Community Health at the University of Maryland School of Public Health, College Park, Maryland (MQW); the Office of Institutional Diversity at the H. Lee Moffitt Cancer Center & Research Institute, Tampa, Florida (BLG); the Department of VA Health Services Research & Development Service at the US Department of Veterans Affairs Healthcare System, Boston, Massachusetts (NRK); the Community Dentistry Section at the University of Puerto Rico School of Dentistry, San Juan, Puerto Rico (CC), and the NYU Langone Medical Center, New York, New York (CS).

Submitted April 9, 2007; accepted November 27, 2007.

Address correspondence to Ralph V. Katz, DMD, MPH, PhD, Department of Epidemiology & Health Promotion, New York University College of Dentistry, 345 E. 24th Street, New York, NY 10012. E-mail: ralph.katz@nyu.edu

Abbreviations used in this paper: TLP = Tuskegee Legacy Project, CSQ = Cancer Screening Questionnaire.

the published articles on factors that affect cancer screening, as reported in two literature reviews, do not assess the influence or impact of "trust" or "fear."[41,42] Rather, the literature on barriers that contribute to the underutilization of cancer screening examinations by minorities has postulated that blacks and Hispanics, as well as those with lower levels of formal education, are less likely to have such examinations due to an array of factors, including limited access to medical care, low income, poor knowledge and attitudes toward the screening process, lack of a regular physician, language barriers, cultural beliefs, and competing demands of day-to-day living.[36,39-49]

A preliminary study by Miller and Hailey[50] in 1994 found that black women experience higher levels of anxiety compared with white women, and the authors suggested that the psychological barriers related to breast cancer screening for black women differ from those for white women. This line of inquiry into the role of emotion and fear as psychological factors affecting cancer screening behaviors has been reported in a recent series of articles by Consedine et al.[51-53] They concluded that the fear factor is one of the key psychological determinants for men related to prostate cancer screening[51] and for women related to breast cancer screening.[52,53] Echoing these findings, a recent report on factors affecting participation in breast cancer studies in the African American community cited "openness to risk" as a key theme.[54]

Underlying the issue of trust or the concept of fear related to any activity for any individual (or within any population subgroup) is the fundamental perception of risks associated with that activity as held by that individual (or within population subgroups). We posit that this concept may underlie both research participation and the use of health services, specifically cancer screening. This report explores this heretofore unevaluated perspective, ie, that of perceived risk, as it relates to participation in biomedical research or in cancer screenings. The specific aim of this analysis was to determine whether there is a difference among blacks, Hispanics, and non-Hispanic whites in their perception of risks associated with participating in either a biomedical study or a cancer screening.

Methods

Subjects and Participants

Data for this report were obtained by the repeated use of several key questions in three separate random-digit-dial (RDD) telephone surveys conducted using the same survey protocol between 1999 and 2003 in seven cities in the United States. Two of the surveys focused on perception of risk associated with participation as a research subject in biomedical studies, and one survey focused on perception of risk associated with having a cancer screening examination.

Survey Instrument

The Tuskegee Legacy Project (TLP) Questionnaire, a 60-item instrument, addresses a range of issues related to the recruitment of minorities into biomedical studies. It was administered via RDD telephone interviews to respondents aged 18 years and older in two separate surveys. In the 4-City TLP Study, the TLP Questionnaire was administered in 1999–2000 in four city areas: Tuskegee, Alabama; Birmingham, Alabama; Hartford, Connecticut; and San Antonio, Texas. The 3-City TLP Study was administered in 2003 in three different cities: New York, New York; Baltimore, Maryland; and San Juan, Puerto Rico. The latter three cities were also used as the sampling base for the administration of the Cancer Screening Questionnaire (CSQ), a derivative questionnaire based on the TLP Questionnaire that focused on a range of issues related to the participation of minorities in cancer screenings. The CSQ was also administered in 2003 but to a separately drawn RDD sample of respondents in the 3-City CSQ Study. Details on the history and development of the TLP Questionnaire and the CSQ, as well as justifications of the methodologic decisions both for the selection of the cities and for the data analysis, are described elsewhere.[15,33-35,43] The 4-City TLP Study was approved by the Institutional Review Boards (IRBs) of the University of Connecticut Health Center and New York University, while the 3-City TLP Study and the 3-City CSQ Study were approved by the IRBs of the University of Puerto Rico and New York University.

4-City TLP Study Sampling Design

The RDD interviews for the 4-City TLP Study were administered by the Survey Research Unit of the University of Alabama at Birmingham (UAB). The target population consisted of noninstitutionalized persons aged 18 years or older living in households with working telephones in the four targeted cities. The RDD sample of households in each of the four cities was based on a simple random sampling of telephone numbers using the three-digit telephone exchanges for those local calling areas with partial screening for nonworking or business numbers. Thirteen interviewers were trained for the survey, using full computer-assisted telephone interviewing (CATI) technology. Unresolved numbers were retired after 20 attempts. Interviewers were trained in calibration sessions and were supervised at all times and randomly electronically monitored a minimum of four times per month.

3-City CSQ and TLP Study Sampling Designs

The RDD telephone interviews for the 3-City TLP Study and the 3-City CSQ Study were administered by ORC Macro Inc (Burlington, Vermont), a US-based international opinion research corporation, using a CATI system for the data collection. The survey sample for this study was drawn from the total noninstitutionalized

adult populations (ages 18 years and over) residing in telephone-equipped dwelling units in three targeted cities. The study provided for a disproportionally allocated, stratified, random-digit sample of telephone-equipped residential households in the three targeted cities, which were sampled independently. The telephone survey followed a 10-attempt dialing protocol in which up to 10 attempts were made unless a final disposition was obtained. Experienced, calibrated supervised personnel conducted the interviews using a CATI software package (CfMC Research Software & Service Bureau, San Francisco, California).

Key Survey Questions

The two key questions that form the focus of this paper were repeated on all three surveys. These questions asked respondents about perceived risks associated with participating as a research subject in a biomedical study or about perceived risks associated with participating in a cancer screening examination. These two key questions on the TLP Questionnaire, as used in both the 4-City and the 3-City surveys, asked, "Do you believe that African Americans in the United States are more likely to be 'taken advantage of' when they become subjects in a medical research practice compared to whites?" and "Do you believe that Hispanics in the United States are more likely to be 'taken advantage of' when they become subjects in a medical research practice compared to whites?" The parallel two questions on the 3-City Cancer Screening Questionnaire asked, "Do you believe that African Americans (or Hispanics) in the United States are less likely to get a 'thorough and careful examination' when they take part in a cancer screening compared to whites?" While the possible responses to each of these questions during questionnaire administration ranged across a 5-point Likert scale (always, most of the time, some of the time, rarely, and never), final analyses used dichotomized responses (always/most of the time vs the other three responses) that had been determined a priori.

The telephone interviewers, regardless of whether they were from the UAB Survey Research Unit (for the 4-City TLP Study) or from ORC Macro Inc (for the 3-City TLP and CSQ Studies), read the same question in the same manner, at the approximate midway point of the questionnaire interview. Moreover, the interviewers in all three surveys read the same choice of possible responses to the respondents: "always," "most of the time," "some of the time," "rarely," and "never" (as well as "don't know"). For purposes of analyses, when dichotomization of the answers were used, "always" and "most of the time" were categorized together and compared with the remaining three valid answer choices ("some of the time," "rarely," and "never").

Response Rate Calculations and Statistical Analyses

The reported response rates are the Council of America Survey Research Organization (CASRO) rates (ie, the percentage of completions of residential households dialed). The reported cooperation rates are the per-

Table 1. — Distribution of the 3,443 Subjects by Age, Sex, Education, and Income Within Ethnic Groups for the 4-City Tuskegee Legacy Project (TLP) Study, the 3-City Tuskegee Legacy Project (TLP) Study, and the 3-City Cancer Screening Questionnaire (CSQ) Study

	4-City TLP Study (1999–2000)			3-City TLP Study (2003)			3-City CSQ Study (2003)		
	Blacks (n = 353)	Whites (n = 623)	Hispanics (n = 157)	Blacks (n = 356)	Whites (n = 493)	Hispanics (n = 313)	Blacks (n = 355)	Whites (n = 482)	Hispanics (n = 311)
Mean age	49.1	53.8	41.5	47.2	48.4	44.3	45.1	47.5	44.0
± SD	± 16.5	± 17.0	± 16.1	± 15.5	± 17.1	± 15.8	± 16.5	± 17.0	± 15.9
Female	47.9%	51.7%	60.5%	67.4%	63.3%	68.4%	72.4%	56.8%	69.5%
Education level									
Less than high school graduate	21.6%	11.8%	14.0%	18.1%	11.8%	21.9%	19.9%	10.6%	19.8%
High school graduate	60.5%	51.3%	61.0%	54.0%	42.2%	41.2%	59.1%	44.0%	43.3%
College graduate or greater	17.9%	36.9%	25.0%	28.0%	45.9%	37.0%	21.0%	45.4%	36.9%
Income level									
< $20,000	42.8%	21.3%	41.7%	33.5%	20.8%	42.3%	41.5%	19.6%	39.2%
$20,000–$74,999	52.1%	58.4%	52.5%	57.8%	56.5%	40.7%	49.5%	59.7%	49.8%
≥ $75,000	5.1%	20.3%	5.8%	8.7%	23.7%	8.0%	9.0%	20.6%	11.0%
Statistically Significant Contrasts* ($P < .05$) Within Each Study:									
Blacks vs whites	Age, education, income			Education, income			Age, sex, education, income		
Blacks vs Hispanics	Age, sex			Age, education			Education		
Hispanics vs whites	Age, sex, education, income			Age, education, income			Age, sex, education, income		

*Using *t* tests for parametric data (age) and chi-square tests for nonparametric data (sex, education level, and income level) with adjustment for multiple comparisons using the Tukey post hoc test criterion.

Total percentages do not equal 100% in all columns due to rounding off of figures.

centage of completed interviews once contact was made with the targeted household subject. These two rates are the universal standard methods of reporting the response and cooperation rates for RDD surveys.[55]

The frequency distribution of all selected variables was examined first. Recoding of categories followed to assure that the frequency distribution was appropriate for the planned statistical analysis. Then, bivariate logistic regression analysis was conducted to examine the relationship between each independent variable and the dependent variable. Finally, the multivariate logistic regression was conducted with race as the independent variable, adjusting for age, sex, education level, income level, and city. All analyses were conducted using either SPSS v14.0 (SPSS Inc, Chicago Illinois) or SAS v9.0 (SAS Institute Inc, Cary, North Carolina) data analysis software.

Results

Demographic Findings

Of the 3,443 respondents in the three surveys in these seven US cities, 30.9% were black, 22.7% were Hispanic, and 46.4% were non-Hispanic white. The 4-City TLP Study, conducted in 1999–2000, included 1,133 respondents with response rates of 70%, 65%, 50%, and 49% across the four cities of Birmingham and Tuskegee in Alabama, San Antonio in Texas, and Hartford in Connecticut, respectively. The 3-City TLP Study conducted in 2003 had 1,162 respondents, with response rates of 52% for San Juan, Puerto Rico, 51% for Baltimore, Maryland, and 44% for New York, New York. The 3-City CSQ Study, with 1,148 respondents, had response rates for these same three cities of 58%, 51%, and 45%, respectively. The Hispanic sample in the 4-City TLP Study was a mix of 75% Mexican Americans and 25% Puerto Rican Americans, while in the two 3-City Studies, the Hispanic sample was all of Puerto Rican descent. The cooperation rate (ie, number of completed interviews and number of initiated interviews) for the 4-City TLP Study was over 90% and the cooperation rate for each of the two 3-City Studies was over 82%. Table 1 shows the age, sex, education, and income distribution of the 3,443 subjects by race for each of the three studies.

Survey Findings

Fig 1 shows the unadjusted percentage of black, Hispanic, and white respondents in each of the three surveys who answered "always" or "most of the time" to the question, "Are African Americans more likely to be 'taken advantage of' when participating in biomedical research [or 'less likely to get a thorough examination' for cancer screenings] than whites?" When asked about being "taken advantage of" when participating in biomedical studies, blacks in the 4-City TLP Study were 8.2 times as likely as whites to say that blacks would "always" or "most of the time" be taken advantage of and 4.3 times as likely in the 3-City TLP Study, while His-

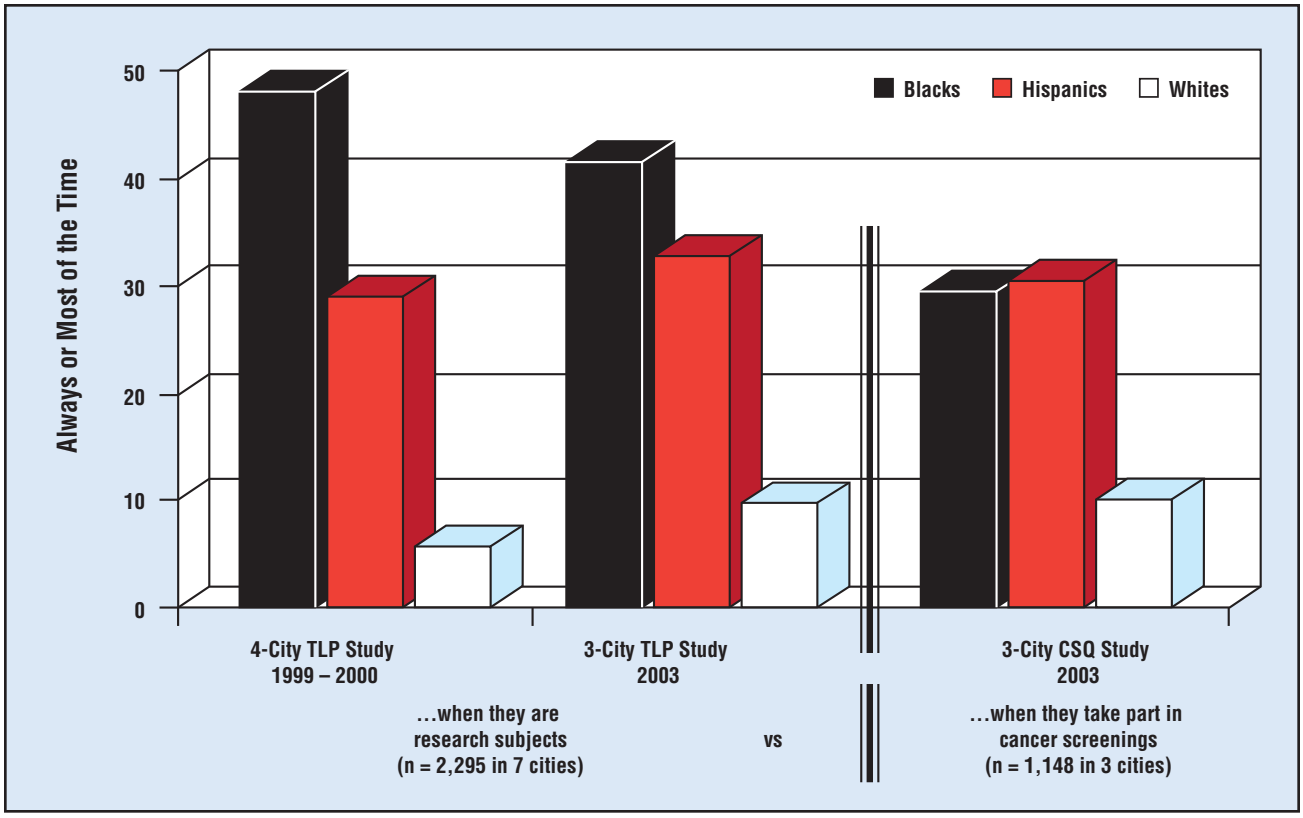


Fig 1. — "Always" and "most of the time" responses to question whether blacks in the United States are more likely to be "taken advantage of" in research studies or "less likely to get a thorough examination" in a cancer screening, compared with whites, based on unadjusted data from three separate studies.

panics were 5.0 and 3.4 as likely as whites to say that African Americans would be taken advantage of "always" or "most of the time," respectively, in the two surveys. When asked about having a cancer screening, blacks and Hispanics were equally likely to state that African Americans would have a less thorough examination "always" or "most of the time" compared with whites (2.9 and 3.0 times as likely, respectively).

Fig 2 shows the unadjusted percentage of black, Hispanic, and white respondents who answered "always" or "most of the time" to the question, "Are Hispanics more likely to be 'taken advantage of' when participating in biomedical research participation [or 'less likely to get a thorough examination' for cancer screenings] than whites?" in each of the three surveys. Regarding being "taken advantage of" when participating in biomedical studies, both blacks and Hispanics were more likely than whites to say that Hispanics would "always" or "most of the time" be "taken advantage of" (9.4 and 6.8 times as likely in the 4-City TLP Study and 3.5 and 3.8 times as likely in the 3-City TLP Study, respectively). Regarding having a cancer screening, blacks and Hispanics were equally as likely to say that Hispanics would have a less thorough examination "always" or "most of the time" compared with whites (3.4 and 3.6 times as likely, respectively).

Multivariate analysis results for the racial/ethnic contrasts for two perceived risks questions for each of the three surveys are shown in Table 2. Overall, chi-square analyses, adjusted for age, sex, education level, income level, and city, revealed that blacks, Hispanics, and whites in a 3-way comparison had statistically significant different response patterns of answering "always" or "most of the time" ($P < .0001$) for both questions. Table 2 shows the odds ratios (ORs) and 95% confidence intervals (CIs) for each specific question for the two-way racial/ethnic contrasts in each of the three surveys, calculated using adjusted logistic regression analyses. When minorities are compared with whites, the ORs range from 3.6 to 14.2 with accompanying CIs that always exclude the value of 1.0, ie, they are all statistically significant. Conversely, when blacks are compared with Hispanics, the ORs range from 0.6 to 1.7 with accompanying CIs that never exclude unity, ie, none of them are statistically significant.

The statistically significant findings from Table 2 are presented graphically in Fig 3 (all ORs were statistically significant except those for the two-way comparisons between blacks and Hispanics). Fig 3 shows the consistency of the OR findings that blacks and Hispanics each believe that minorities, compared with whites, are much more likely to be "taken advantage of" in biomedical studies and much less likely to get a "thorough and careful examination" in a cancer screening across the two key questions (Q31 and Q32) within each of the three studies.

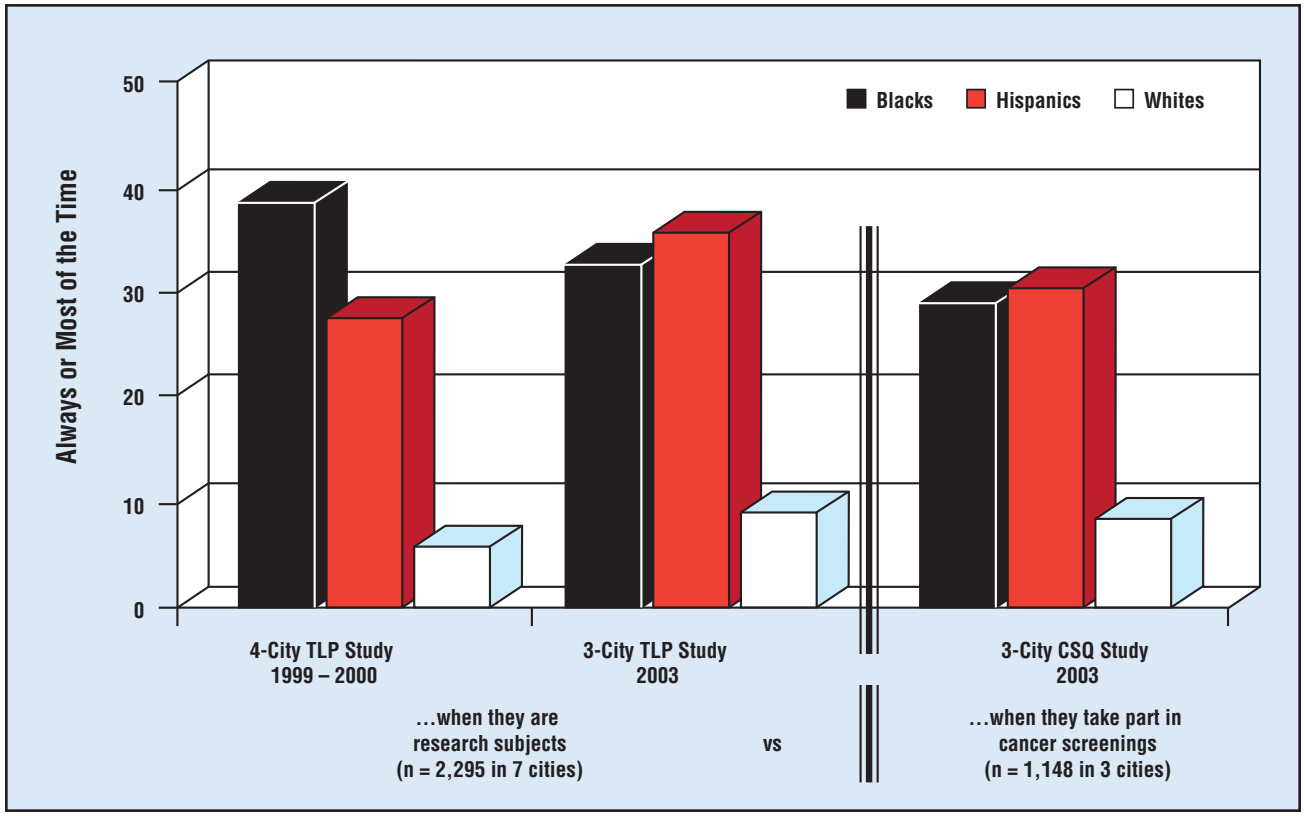


Fig 2. — "Always" and "most of the time" responses to question whether Hispanics in the United States are more likely to be "taken advantage of" in research studies or "less likely to get a thorough examination" in a cancer screening, compared with whites, based on unadjusted data from three separate studies.

Discussion

These data show overall similarities in both the direction and magnitude of the differences in "perceived risk" for both the blacks and Hispanics, compared with the perception of whites, as related to participation in both biomedical studies and cancer screenings. In addition, within the studies that focused on perceived risks of participating in biomedical research (ie, the 1999–2000 4-City TLP Study and the 3-City TLP Study 2003), the difference in perceived risks between each minori-

Table 2. — Summary of Adjusted* Logistic Regression Analysis of Q31-32 for the 4-City Tuskegee Legacy Project (TLP) Study (1999–2000), the 3-City TLP Study (2003) and the 3-City Cancer Screening Questionnaire (CSQ) Study (2003)

	Q31: Are blacks more likely to be "taken advantage of" than whites … … when becoming subjects in a biomedical study?		Are blacks less likely to get a thorough examination than whites … … when participating in a cancer screening?
Racial/Ethnic Contrasts	**4-City TLP Study (1999–2000) OR (95% CI)**	**3-City TLP Study (2003) OR (95% CI)**	**3-City CSQ Study (2003) OR (95% CI)**
Blacks vs Whites[a]	**14.2 (8.9–22.6)**	**5.6 (2.9–10.9)**	**4.1 (2.1–8.1)**
Hispanics vs Whites[a]	**7.3 (3.2–16.6)**	**3.9 (1.5–10.5)**	**3.6 (1.7–7.9)**
Blacks vs Hispanics[b]	1.7 (0.8–3.4)	1.2 (0.6–2.6)	1.1 (0.5–2.4)
	Q32: Are Hispanics more likely to be "taken advantage of" than whites … … when becoming subjects in a biomedical study?		**Are Hispanics less likely to get a thorough examination than whites … … when participating in a cancer screening?**
Racial/Ethnic Contrasts	**4-City TLP Study (1999–2000) OR (95% CI)**	**3-City TLP Study (2003) OR (95% CI)**	**3-City CSQ Study (2003) OR (95% CI)**
Blacks vs Whites[a]	**12.1 (7.2–20.4)**	**3.9 (2.0–7.7)**	**3.8 (1.9–7.5)**
Hispanics vs Whites[a]	**7.4 (3.1–17.3)**	**5.5 (2.2–14.1)**	**4.5 (2.1–9.4)**
Blacks vs Hispanics[b]	1.2 (0.6–2.4)	0.6 (0.3–1.2)	1.0 (0.5–2.1)

* adjusted for age, sex, education, income, and city
[a] referent group: Whites
[b] referent group: Hispanics
All statistically significant findings are in bold type.
The adjusted 3-way analyses by race/ethnic group for both questions (Q31 and Q32) were statistically significant (χ^2, $P < .0001$).

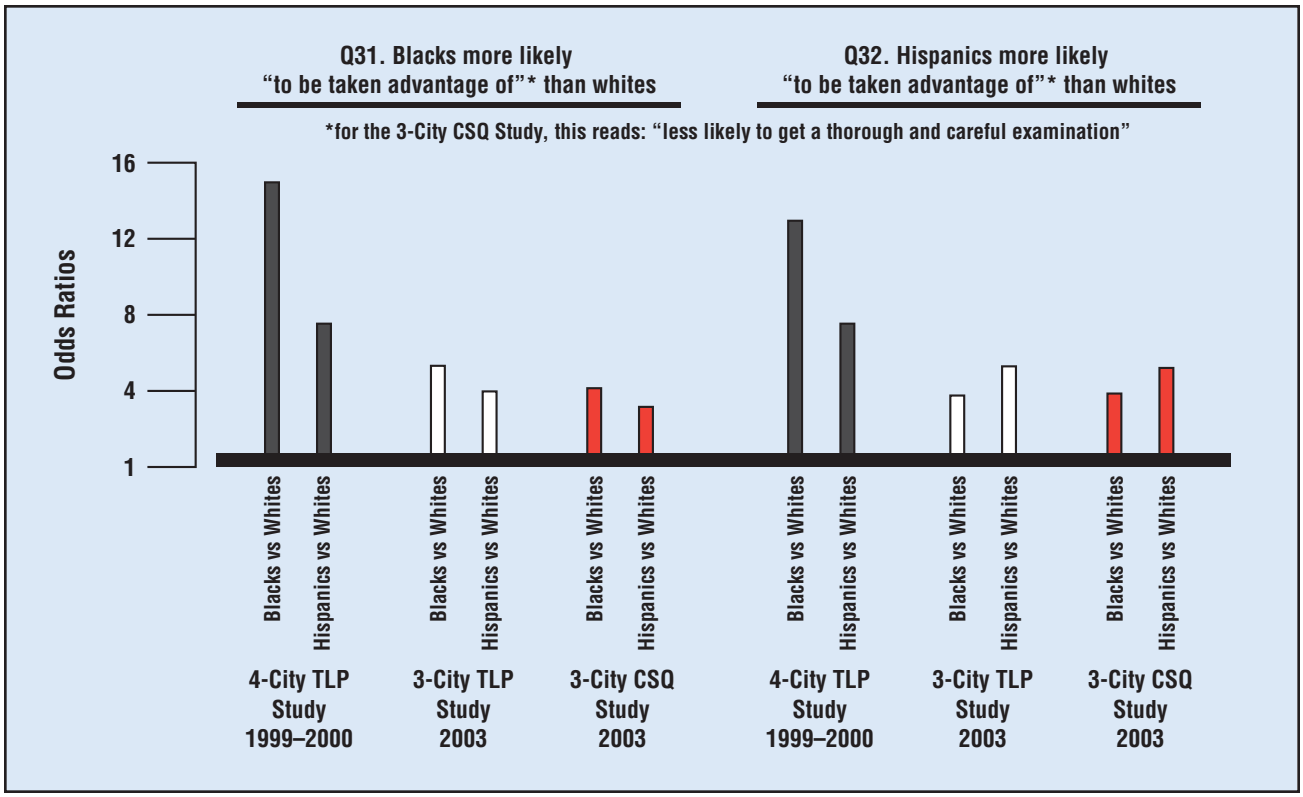


Fig 3. — Statistically significant odds ratios (ORs) for two-way racial/ethnic contrasts for the perceived risks when participating in biomedical research studies and cancer screenings in three separate studies, adjusted for age, sex, education, income, and city.

ty group vs whites was markedly and consistently higher in the earlier study. This observed 2.5-fold lowering in the odds of perceived risks for participating as research subjects for blacks vs whites in the 1999–2000 4-City TLP Study (OR = 14.2) compared with the 2003 3-City TLP Study (OR = 5.6) might be due to the temporal proximity of that 1999–2000 4-City TLP Study to the well-publicized Presidential Apology for the United States Public Health Service Tuskegee Syphilis Study as made by President Clinton in 1997.[34] Alternatively, this observation might be due to regional differences in the two study samples, as two-thirds of the blacks in the 4-City TLP Study came from US cities in the deep South (ie, Alabama) whereas the southernmost city in the 3-City TLP Study was Baltimore, in the Middle Atlantic state of Maryland, or might reflect the influence of both of these factors.

This dichotomous minority vs white viewpoint of "perceived risks" regarding these two types of events was most exaggerated in the comparison on research subject participation between the earlier 4-City TLP Study (1999–2000) and the later 3-City TLP Study (2003). Even putting that largest discrepancy aside, the data from the two 2003 studies show that blacks and Hispanics perceive themselves as a group to be a greater risk than whites when they participate as research subjects (ie, 4-fold or greater odds of "being taken advantage of" always or most of the time). Interestingly, this viewpoint is also held regarding participation in cancer screenings (and at the same magnitude of perceived risk, ie, about a 4-fold greater odds for blacks and Hispanics).

These data, specific to perceived risks associated with either participation in cancer screenings or biomedical research, appear to echo the well-known and broader society issue of differing perceptions by blacks and whites on life in the United States.[16-18,34] The perception of risks of life in the United States appears to be highly dependent on whether one is viewing US life through the prism provided by the perceptions of the black or Hispanic community or the prism provided by perceptions of the white community. One of the most recent well-publicized examples of the existence of these two prisms by which life in the United States is perceived was provided by the worldwide coverage of the O.J. Simpson murder trial. Following the verdict of "not guilty," the Gallup Poll of October 5–7, 1995, reported that the majority (78%) of the black community agreed with the "not guilty" decision, while a minority (42%) of the white community agreed with that verdict.[56] This vast divergence in the expressed opinion of the black community and the white community of the United States as to the correctness of that verdict speaks directly to the widespread use of one or the other of these two prisms of life in the United States. The data from these three surveys reported here on perceived risks associated with participating in cancer screenings and biomedical studies demonstrate how deeply these two prisms reflect into other — if not all — aspects of views on life in the United States.

Conclusions

While volunteering to be a subject in a research study inherently and overtly carries clear "risks of participation" for any subject, it is more surprising that this concept of "being taken advantage of" carries over so equally into the arena of standard health care procedures, such as routine cancer screening examinations. The fact that blacks and Hispanics perceive equally high levels of risk for participating in cancer screening examinations as they do for volunteering to become research subjects in biomedical studies provides a strong message to health care providers, organizations, and systems about the need to overtly address this critical health disparities issue. Given that most cancer incidence rates disproportionately affect minorities in the United States and that most cancer mortality rates also disproportionately burden minorities, there is a clear need for all service components of the US health care system to focus on improving their cancer screening actions and cancer screening messages to the minority communities that they serve.

Disclosures

The three studies that provided these data, as well as the data analyses, were supported by two National Institute of Dental and Craniofacial Research/National Institutes of Health grants U54 DE 14257 and P50 DE10592.

No significant relationship exists between the authors and the companies/organizations whose products or services may be referenced in this article.

References

1. Healthy People 2010: Understanding and Improving Health. 2nd ed. Goal 2: Eliminate Health Disparities. http://www.healthypeople.gov/Document/html/uih/uih_2.htm#goals. Accessed June 3, 2008.
2. National Institutes of Heath (NIH) Guidelines on Inclusion of Women and Minorities as Subjects in Clinical Research. *NIH Guide.* 1994;23:(11). http://www.grants.nih.gov/grants/guide/notice-files/not94-100.html. Accessed June 3, 2008.
3. Smith RA, Cokkinides V, Eyre HJ. American Cancer Society guidelines for the early detection of cancer, 2006. *CA Cancer J Clin.* 2006;56(1):11-25; quiz 49-50.
4. Jemal A, Siegel R, Ward E, et al. Cancer statistics, 2006. *CA Cancer J Clin.* 2006;56(2):106-130.
5. Kerr AR, Changrani JG, Gany FM, et al. An academic dental center grapples with oral cancer disparities: current collaboration and future opportunities. *J Dent Educ.* 2004;68(5):531-541.
6. Sankaranarayanan R, Ramadas K, Thomas G, et al. Effect of screening on oral cancer mortality in Kerala, India: a cluster-randomised controlled trial. *Lancet.* 2005;365(9475):1927-1933.
7. Hunninghake DB, Darby CA, Probstfield JL. Recruitment experience in clinical trials: literature summary and annotated bibliography. *Control Clin Trials.* 1987;8(4 suppl):6S-30S.
8. Swanson GM, Ward AJ. Recruiting minorities into clinical trials: toward a participant-friendly system. *J Natl Cancer Inst.* 1995;87(23):1747-1759.
9. Gamble VN. A legacy of distrust: African Americans and medical research. *Am J Prev Med.* 1993;9(6 suppl):35-38.
10. Caplan AL. Twenty years after: the legacy of the Tuskegee Syphilis Study. When evil intrudes. *Hastings Cent Rep.* 1992;22(6):29-32.
11. Corbie-Smith GM. Minority recruitment and participation in health research. *N C Med J.* 2004;65(6):385-387.
12. Rathore SS, Krumholz HM. Race, ethnic group, and clinical research. *BMJ.* 2003;327(7418):763-764.

13. Boulware LE, Cooper LA, Ratner LE, et al. Race and trust in the health care system. *Public Health Rep*. 2003;118(4):358-365.
14. Brandon DT, Isaac LA, LaVeist TA. The legacy of Tuskegee and trust in medical care: is Tuskegee responsible for race differences in mistrust of medical care? *J Natl Med Assoc*. 2005;97(7):951-956.
15. Katz RV, Kegeles SS, Green BL, et al. The Tuskegee Legacy Project: history, preliminary scientific findings and unanticipated societal benefits. *Dent Clin North Am*. 2003;47(1):1-19.
16. White RM. Misinformation and misbeliefs in the Tuskegee Study of Untreated Syphilis fuel mistrust in the healthcare system. *J Natl Med Assoc*. 2005;97(11):1566-1573.
17. Fairchild AL, Bayer R. Uses and abuses of Tuskegee. *Science*. 1999; 284(5416):919-921.
18. Washington HA. *Medical Apartheid: The Dark History of Medical Experimentation on Black Americans from Colonial Times to the Present*. New York, NY: Doubleday Press; 2006.
19. Corbie-Smith G, Thomas SB, Williams MV, et al. Attitudes and beliefs of African Americans toward participation in medical research. *J Gen Intern Med*. 1999;14(9):537-546.
20. Green BL, Partridge EE, Fouad MN, et al. African-American attitudes regarding cancer clinical trials and research studies: results from focus group methodology. *Ethn Dis*. 2000;10(1):76-86.
21. Freimuth VS, Quinn SC, Thomas SB, et al. African Americans' views on research and the Tuskegee Syphilis Study. *Soc Sci Med*. 2001;52(5): 797-808.
22. Bates BR, Harris TM. The Tuskegee Study of Untreated Syphilis and public perceptions of biomedical research: a focus group study. *J Natl Med Assoc*. 2004;96(8):1051-1064.
23. Sengupta S, Strauss RP, DeVellis R, et al. Factors affecting African-American participation in AIDS research. *J Acquir Immune Defic Syndr*. 2000; 24(3):275-284.
24. Brown DR, Topcu M. Willingness to participate in clinical treatment research among older African Americans and Whites. *Gerontologist*. 2003; 43(1):62-72.
25. Green BL, Maisiak R, Wang MQ, et al. Participation in health education, health promotion, and health research by African Americans: effects of the Tuskegee Syphilis Experiment. *J Health Educ*. 1997;28(4):196-201.
26. Shavers VL, Lynch CF, Burmeister LF. Knowledge of the Tuskegee Study and its impact on the willingness to participate in medical research studies. *J Natl Med Assoc*. 2000;92(12):563-572.
27. McCallum JM, Arekere DM, Green BL, et al. Awareness and knowledge of the U.S. Public Health Service Syphilis Study at Tuskegee: implications for biomedical research. *J Health Care Poor Underserved*. 2006;17(4): 716-733.
28. Kressin NR, Meterko M, Wilson NJ. Racial disparities in participation in biomedical research. *J Natl Med Assoc*. 2000;92(2):62-69.
29. Shavers-Hornaday VL, Lynch CF, Burmeister LF, et al. Why are African Americans under-represented in medical research studies? Impediments to participation. *Ethn Health*. 1997;2(1-2):31-45.
30. Chávez V, Duran B, Baker QE, et al. The dance of race and privilege in community-based participatory research. In: Minkler M, Wallerstein N, eds. *Community Based Participatory Research for Health*. San Francisco, CA: Jossey Bass; 2002:81-97.
31. Hatch J, Moss N, Saran A, et al. Community research: partnership in black communities. *Am J Prev Med*. 1993;9(6 suppl):27-31.
32. Matthews AK, Sellergren SA, Manfredi C, et al. Factors influencing medical information seeking among African American cancer patients. *J Health Commun*. 2002;7(3):205-219.
33. Katz RV, Kegeles SS, Kressin NR, et al. The Tuskegee Legacy Project: willingness of minorities to participate in biomedical research. *J Health Care Poor Underserved*. 2006;17(4):698-715.
34. Katz RV, Kegeles SS, Kressin NR, et al. Awareness of the Tuskegee Syphilis Study and the US presidential apology and their influence on minority participation in biomedical research. *Am J Public Health*. 2008;98(6): 1137-1142. Epub 2007 Sep 27.
35. Katz, RV, Green, BL, Kressin, NR, et al. Willingness of minorities to participate in biomedical studies: confirmatory findings from a follow-up study using the Tuskegee Legacy Project Questionnaire. *J Natl Med Assoc*. 2007;99(9):1052-1060. Erratum in: *J Natl Med Assoc*. 2007;99(11):1216.
36. Abraído-Lanza AF, Chao MT, Gammon MD. Breast and cervical cancer screening among Latinas and non-Latina whites. *Am J Public Health*. 2004;94(8):1393-1398.
37. Swan J, Breen N, Coates RJ, et al. Progress in cancer screening practices in the United States: results from the 2000 National Health Interview Survey. *Cancer*. 2003;97(6):1528-1540.
38. Zambrana RE, Breen N, Fox SA, et al. Use of cancer screening practices by Hispanic women: analyses by subgroup. *Prev Med*. 1999;29(6, pt 1): 466-477.
39. Luquis RR, Villanueva IJ. Knowledge, attitudes, and perceptions about breast cancer and breast cancer screening among Hispanic women residing in South Central Pennsylvania. *J Community Health*. 2006;31(1):25-42.
40. Abraído-Lanza AF, Chao MT, Gates CY. Acculturation and cancer-screening among Latinas: results from the National Health Interview Survey. *Ann Behav Med*. 2005;29(1):22-28.
41. Vernon SW, Laville EA, Jackson GL. Participation in breast cancer screening programs: a review. *Soc Sci Med*. 1990;30(10):1107-1118.
42. Austin L, Ahmad F, McNally MJ, et al. Breast and cervical cancer screening in Hispanic women: a literature review using the health belief model. *Womens Health Issues*. 2002;12(3):122-128.
43. Claudio C, Katz RV, Green BL, et al. Cancer screening participation: comparative willingness of San Juan Puerto Ricans vs New York City Puerto Ricans. *J Natl Med Assoc*. 2007;99(5):542-549.
44. Kunkel EJ, Meyer B, Daskalakis C, et al. Behaviors used by men to protect themselves against prostate cancer. *Cancer Epidemiol Biomarkers Prev*. 2004;13(1):78-86.
45. Odedina FT, Scrivens J, Emanuel A, et al. A focus group study of factors influencing African-American men's prostate cancer screening behavior. *J Natl Med Assoc*. 2004;96(6):780-788.
46. Magai C, Consedine NS, Conway F, et al. Diversity matters: unique populations of older women and breast cancer screening. *Cancer*. 2004; 100(11):2300-2307.
47. Bloom J, Hayes W, Saunders F, et al. Cancer awareness and secondary prevention practices in black Americans: implications for intervention. *Fam Commun Health*. 1987;10:19-30.
48. Friedman LC, Webb JA, Weinberg AD, et al. Breast cancer screening: racial/ethnic differences in behaviors and beliefs. *J Cancer Educ*. 1995; 10(4):213-216.
49. Aro AR, de Koning HJ, Absetz P, et al. Two distinct groups of non-attenders in an organized mammography screening program. *Breast Cancer Res Treat*. 2001;70(2):145-153.
50. Miller LY, Hailey BJ. Cancer anxiety and breast cancer screening in African-American women: a preliminary study. *Womens Health Issues*. 1994; 4(3):170-174.
51. Consedine NS, Morganstern AH, Kudadjie-Gyamfi E, et al. Prostate cancer screening behavior in men from seven ethnic groups: the fear factor. *Cancer Epidemiol Biomarkers Prev*. 2006;15(2):228-237.
52. Consedine NS, Magai C, Neugut AI. The contribution of emotional characteristics to breast cancer screening among women from six ethnic groups. *Prev Med*. 2004;38(1):64-77.
53. Consedine NS, Magai C, Krivoshekova YS, et al. Fear, anxiety, worry, and breast cancer screening behavior: a critical review. *Cancer Epidemiol Biomarkers Prev*. 2004;13(4):1-10.
54. Linden HM, Reisch LM, Hart A Jr, et al. Attitudes toward participation in breast cancer randomized clinical trials in the African American community: a focus group study. *Cancer Nurs*. 2007;30(4):261-269.
55. Council of American Survey Research Organizations (CASRO). *Report of the CASRO Completion Rates Task Force*. New York, NY: Audits and Surveys Co Inc; 1982.
56. Gallup Poll, October 5-7, 1995. http://www.galluppoll.com. Accessed June 9, 2008.

Teaching and Learning in Medicine, 20(2), 143–150

ISSN: 1040-1334 print / 1532-8015 online
DOI: 10.1080/10401330801991667

Routledge
Taylor & Francis Group

Disagreement Between Students and Preceptors Regarding the Value of Teaching Behaviors for Ambulatory Care Settings

Walter N. Kernan
Department of Internal Medicine, Yale University School of Medicine, New Haven, Connecticut, USA

Warren Hershman
Department of Internal Medicine, Boston University School of Medicine, Boston, Massachusetts, USA

Eric J. Alper
Department of Medicine, University of Massachusetts School of Medicine, Worcester, Massachusetts, USA

Mary Y. Lee
Department of Internal Medicine, Tufts University School of Medicine, Boston, Massachusetts, USA

Catherine M. Viscoli
Department of Internal Medicine, Yale University School of Medicine, New Haven, Connecticut, USA

John R. Perry
University of Cambridge, Cambridge, United Kingdom

Patrick G. O'Connor
Department of Internal Medicine, Yale University School of Medicine, New Haven, Connecticut, USA

Background: **Medical students and preceptors commonly disagree on methods of clinical instruction in ambulatory care, although the extent of the problem is not documented.** ***Purpose:*** **The purpose is to identify disagreement and concordance between students and preceptors for teaching behaviors in ambulatory care.** ***Methods:*** **We surveyed students and preceptors at 4 U.S. schools. Respondents rated 58 behaviors on two scales. Disagreement was recognized when the percentage of students and preceptors who recommended a behavior *and* rated it important differed by over 15% ($p < .01$).** ***Results:*** **Disagreement was identified for 8 behaviors (14%). Six were valued less by students, including "watch the student perform critical tasks in history taking and other communication" (59% compared with 82%). Two behaviors were valued more by students, including "delegate responsibility to the student for the wrap up discussion with the patient" (82% compared with 61%).** ***Conclusions:*** **Students and preceptors disagree regarding the value of a minority of teaching behaviors. Because some are potentially important, however, early negotiation regarding their use may enhance teaching effectiveness and mutual satisfaction with learning.**

In the past 15 years, medical schools throughout the United States have instituted clerkships in ambulatory care internal medicine and other office-based specialties.[1–3] Designed to ensure that students acquire fundamental skills needed for further specialty training, these clerkships involve a complex interaction between student and teacher that is often described as an apprenticeship.[4] The student is acquiring the skills of a physician by doing the work of patient care. The teacher strives to support the student's autonomy while helping him or her grow in skill, knowledge, and professional attitude. All of this occurs in the context of a patient who needs personal, high-quality care for which the teacher is ultimately responsible.

Realizing that the success of this apprenticeship is highly dependent on the teaching skill of the teacher, medical schools have invested considerably in local and regional faculty development.[5–9] These efforts have disseminated

Final revision received July 7, 2007.
Supported by an anonymous gift by an individual to the Department of Medicine at Yale. We acknowledge the many medical students and preceptors who gave generously of their time to make this research possible.
Correspondence may be sent to Walter N. Kernan, MD, IRIS Coordinating Center, Suite 515, 2 Church Street South, New Haven, CT 06519, USA. E-mail: Walter.Kernan@Yale.edu

models of learner-centered education that emphasize personalized instruction and experiences that account for individual learning needs.

Recommended strategies for learner-centered education in ambulatory care settings emphasize greater flexibility in teaching content than in teaching technique. For content, preceptors are advised to identify each student's personal learning goals and capabilities so that experiences can be customized to catalyze continued growth. For teaching technique, a relatively limited range of behaviors is typically prescribed and less emphasis is placed on discussion between student and teacher that might lead to adjustment or customization of instruction. The potential value of improved communication is apparent when students and teachers disagree on the appropriateness or effectiveness of specific teaching behaviors. Disagreements may include student objection to the use of widely recommended behaviors or nonuse of others. Clerkship directors commonly learn about these disagreements when students express dissatisfaction with teaching encounters.

Our study was designed to estimate the frequency of disagreement between students and preceptors for the average value of teaching behaviors available for use in ambulatory care internal medicine. Our hypothesis was that disagreement would be common. Recognition and management of disagreement might provide a basis for enhancing the learner-centered approach to clinical teaching and improving ambulatory care education. A secondary aim of this study was to identify teaching behaviors that are valued by both students and preceptors.

METHODS

Identification of Teaching Behaviors

We used focus group and survey strategies to identify a comprehensive list of teaching behaviors valued by 3rd-year medical students or faculty preceptors participating in ambulatory care clerkships in internal medicine. Behaviors valued by students were identified during research we conducted from 1996 to 1998 at three New England schools.[10] This research yielded 94 behaviors, of which 51 met the prespecified criterion for "valued" (i.e., they were recommended and rated important by $\geq$75% of students). The criterion was developed by consensus among the investigators.

To identify teaching behaviors valued by preceptors, we conducted seven focus groups at three schools (University of Massachusetts, Yale, Boston University) from 1997 to 2000. Participants were selected by the local investigator based on their teaching experience and their effectiveness as clinical teachers. The seven focus groups ranged in size from 2 to 5 preceptors and were conducted according to customary methods.[11] Twenty-two preceptors participated. New groups were assembled until no new teaching behaviors were forthcoming. All discussions were tape-recorded and transcribed. Transcripts from focus groups and minutes from evaluation sessions were read by all five investigators who agreed on a final list of specific behaviors. Of these, 21 had not emerged from the previous research on students.

Ascertainment of Student and Preceptor Preferences for the Behaviors

To ascertain and compare student and preceptor preferences for the teaching behaviors, we created a survey for which respondents rated 58 teaching behaviors according to two scales. The behaviors included 37 selected from the original research on students' preferences[10] and all 21 behaviors from the preceptor focus groups, which had not been identified in the student survey. We selected only 37 behaviors from the original research to limit the length of the new survey. Selection criteria were nonredundancy, specificity, and a mix of valued and not-valued behaviors. For the first scale, respondents were asked, "Do you recommend preceptors use the behavior?" Five response options were *yes, strongly*; *yes, somewhat*; *not sure*; *no, somewhat*; and *no, strongly*. For the second scale, respondents were asked, "How important is the behavior to your learning?" Five response options were *extremely important*, *very important*, *somewhat important*, *not very important*, and *not at all important*.

The population for the survey comprised 50 students and 50 general medicine preceptors at each of the four participating schools. Investigators at each institution invited all students in the last week of consecutive ambulatory care internal medicine clerkships to complete the survey. Only students who agreed to participate were given a survey, but usually all students agreed. Preceptors were recruited from a random sample at each participating institution. A survey instrument was mailed to each randomly identified preceptor. The sample included 39% (200/517) of all eligible students at the four institutions and approximately 53% (200/374) of active general medicine preceptors.

Analysis

For each of the 58 teaching behaviors, we calculated separately the proportions of students and preceptors that valued a behavior. A behavior was defined as valued if it was recommended strongly or somewhat *and* rated extremely or very important. We chose this method of analysis because it was the method we used in our original research and permitted a direct comparison with the earlier findings.[10]

According to criteria developed by consensus among the investigators before data analysis, disagreement was recognized when the absolute difference between proportions of students and preceptors that valued a behavior exceeded 15% at a significance level of $p < .01$ (using chi-square test of differences in proportions). A behavior was classified as valued by students or preceptors if it was valued by 75% or more of group respondents.

TABLE 1
Selected features of students and preceptors

Feature	% of students[a]	% of preceptors[b]	p
Age (M years $\pm$ SD)	27 $\pm$ 3	45 $\pm$ 8	< .001
Female Sex	54	37	.003
Practice Type[c]			
Solo	9	8	.04
Group	29	36	
Hospital	43	32	
Community Health Center	10	20	
Staff Model HMO	6	2	
Other	3	2	
Ethnicity			
White	68	82	.06
Black	5	1	
Hispanic	5	1	
Asian	19	12	
Other	3	1	
Faculty Appointment			
Part-Time	n/a	44	
Full-Time	n/a	40	
None	n/a	16	
Years Precepting (M)	n/a	9 $\pm$ 6	
Faculty Development			
Yes	n/a	74	
No	n/a	26	
Stipend for Teaching			
Yes	n/a	30	
No	n/a	70	

Note: n/a = not applicable.
[a] N = 163. [b] N = 138. [c] For students, practice type refers to the main teaching site to which he or she was assigned. Practice type was missing for one faculty member.

RESULTS

Study Population

Among 200 students who were invited to complete the survey, 163 (82%) responded. Among 200 teachers, 138 (69%) responded. Selected features of respondents are described in Table 1. Compared with students, preceptors were older, more likely to be male, and more likely to be white.

Behaviors Valued Differently by Students and Preceptors

Among the 58 rated behaviors, significant disagreement was observed for 8 (14%) (Table 2). Six were more commonly valued by preceptors compared with students, including watching students perform critical tasks in history taking and other communications with patients, which was valued by 58.3% of students compared with 84.7% of preceptors and associated with the greatest discordance (absolute difference = 26.4%). The other six behaviors more valued by preceptors involved etiquette that may affect the learning environment (i.e., introduce the student to patients using the student's correct name), student–teacher communication required to oversee the student's experience (i.e., periodically inquire about how the experience could be adjusted to better suit the student's needs, periodically ask the student if his or her personal learning goals are being met), orchestrating the student–patient encounter (i.e., ask the student to present the history and physical examination in front of the patient), and clinical skills instruction related to efficiency (i.e., counsel the student on conducting a problem-focused patient encounter).

Two behaviors were more valued by students compared with preceptors (Table 2, Behaviors 2.7, 2.8); both involved expanding student involvement in patient care (i.e., ask the students to do minor procedures and delegate responsibility to the student for the wrap-up discussion with the patient). A third item (Table 4, Behavior 4.8) also involving expanding responsibility and was valued more by students compared with preceptors (i.e., delegate responsibility to the student for ascertaining and interpreting test results) but the difference did not quite reach the 15% criterion (valued by 82.5% of students compared with 67.6% of preceptors, difference = 14.9, p = .003).

Behaviors Valued by Both Students and Preceptors

Among the 58 behaviors examined, 32 were recommended strongly or somewhat *and* rated extremely or very important by 75% or more of both students and preceptors (Table 3). For each of the 32, the difference in the percentage of students and preceptors who valued them was small and did not reach our criterion for disagreement (i.e., >15%).

Twelve behaviors were valued by over 90% of respondents in both groups. Most (8) of these involved the two domains of teaching clinical skills and feedback. From the domain of teaching clinical skills, the most highly rated behavior involved challenging the student to explain choices he or she makes regarding diagnostic strategies or therapeutics (Table 3, Behavior 3.9), followed closely by guiding the student in devising a plan of care and caring for the patient while avoiding replacing the student or just telling the student what to do (Table 3, Behavior 3.10), assuring the student regularly interviews and examines patients on his or her own (Table 3, Behavior 3.11), and asking for the student's assessment and plan before giving one's own formulation (Table 3, Behavior 3.12). From the domain of feedback, three of four behaviors valued by more than 90% of students and preceptors were very similar and involved following honest criticism with provision of specific help toward improvement (Table 3, Behaviors 3.27–3.29).

The remaining behaviors valued by both students and preceptors involve most domains of clinical teaching except orientation to the rotation (Table 3).

TABLE 2
Eight teaching behaviors valued differently by students and preceptors, listed according the magnitude of the difference in the percentage of students and preceptors who valued each

No.	Behavior	% Respondents Valuing the Behavior			
		Students[a]	Preceptors[b]	Difference	*p*
2.1	Regularly watch the student perform critical tasks in history taking and other patient communications.[c]	58.3	84.7	–26.4	.000
2.2	Early in the rotation, counsel the student on conducting a problem-focused patient encounter.[c]	67.3	89.1	−21.8	.000
2.3	Introduce the student to patients using the student's correct name.	45.4	67.2	−21.8	.000
2.4	Periodically inquire about how the experience could be adjusted to better suit the student's needs.[d]	61.7	82.2	−20.5	.000
2.5	Periodically ask the student if his or her personal learning goals are being met.[c]	64.2	84.4	−20.2	.000
2.6	For most patients, ask the student to present the history and physical examination in front of the patient.[e]	12.5	27.8	−15.3	.001
2.7	Delegate responsibility to the student for the wrap-up discussion with the patient (for explaining the diagnosis and treatment, etc.).[d]	78.9	59.3	19.6	.000
2.8	Ask the student to do minor procedures, such as injections, tuberculin skin testing, and electrocardiogram interpretation.[d]	89.6	70.8	18.8	.000

[a] $N = 163$. [b] $N = 138$. [c] Item was identified only during faculty focus groups. [d] Identical or very similar items were identified as valued by students in a previous survey.[10] [e] Identical or very similar to items that were identified as not valued by students in a previous survey.[10]

Behaviors Not Valued by Both Students and Preceptors

Among the 58 behaviors examined, 26 were valued by students alone ($n = 3$), preceptors alone ($n = 9$), or neither ($n = 14$). These 26 included all 8 behaviors for which the proportion of students and preceptors who valued the behavior differed by more than 15% (Table 2) and 18 for which the difference was smaller (Table 4). The least valued behavior was questioning students about medical knowledge in front of patients (Table 4, Behavior 4.16).

DISCUSSION

Our findings identify a large number of specific teaching behaviors valued by both students and preceptors, and a smaller but significant number of behaviors about which they disagree. Among the eight behaviors for which we observed disagreement, six were more highly valued by preceptors and involved techniques to enhance student efficiency or monitor student progress. Two were more highly valued by students compared with preceptors and involved giving students broader responsibilities in patient care, including minor procedures and visit closures.

As in our previous work,[10] students expressed a distinct lack of enthusiasm for presentations in the examination room (Table 2, Behavior 2.6). Although our data do not provide a direct explanation for this aversion, students apparently do not like being questioned about their medical knowledge in front of patients (Table 4, Behavior 4.16). Other investigators have found that students prefer to present outside the examination room because they believe there may be more time for teaching and questions, they are uncomfortable presenting in the room, they believe patients are uncomfortable, or they dislike editing their discourse for patients.[12]

To our knowledge only one other study has examined the phenomenon of disagreement for specific teaching behaviors between groups of learners and teachers in clinical medicine.[13] Investigators at the Mayo Clinic in Scottsdale, Arizona, asked 179 residents and 117 faculty members in eight U.S. family medicine residency programs to review a list of 15 teaching attributes before indicating the three most and least important. Disagreement was recognized when the *p*value was less than .05 for the difference in proportion of residents and faculty members who ranked a behavior among the "top three." Among the four behaviors (27%) meeting the criteria for disagreement, residents were more likely to value a preceptor who supported their autonomy and less likely to value role modeling.

How preceptors handle disagreement may affect student satisfaction with ambulatory education and their mastery of ambulatory care skills. Based on our findings, preceptors should anticipate that students will object to some behaviors and welcome others. Advance discussion about all potential behaviors and expectations may foster a more collaborative learning environment. For example, a preceptor who stays in the examination room to watch a student communicate with the patient may

TABLE 3
Thirty-two teaching behaviors valued by both students and preceptors, ranked within domains according to student responses

No.	Behavior	% Respondents Valuing the Behavior			
		Students[a]	Preceptors[b]	Absolute Difference	*p*
Domain: Orientation to the Rotation					
	None				
Domain: Creating a Favorable Learning Environment					
3.1	Encourage students to ask questions throughout the rotation.[c]	93.9	97.8	−3.9	.095
3.2	Encourage questions and respond to them tactfully.[c]	92.6	98.5	−5.9	.016
3.3	Initiate teaching discussions.[c]	91.4	86.9	4.5	.204
Domain: Overseeing the Student's Experience					
3.4	Ask the student if there are aspects of the physical examination he or she wants to work on and then provide help.[d]	92.0	88.9	3.1	.365
3.5	Look out for learning opportunities for the student. For example, if a patient needs a procedure, have the student do it.[c]	90.8	84.3	6.5	.089
3.6	Enable the student to see a mix of acute visit patients and non-acute visit patients.[c]	88.3	84.3	4.0	.323
3.7	Early in the rotation, ask the student to identify skills he or she wants to develop.[c]	79.8	75.7	4.1	.404
Domain: Orchestrating the Student–Patient Interaction					
3.8	If the student presents the history and physical in front of the patient, provide the student an opportunity to also talk to the preceptor away from the patient.[e]	78.5	84.3	−5.8	.203
Domain: Teaching Clinical Skills					
3.9	Challenge the student to explain choices he or she makes regarding diagnostic strategies or therapeutics.	97.5	99.3	−1.8	.246
3.10	Guide the student in devising a plan of care and caring for the patient; avoid replacing the student or just telling the student what to do.[c]	96.9	94.9	2.0	.369
3.11	Assure the student regularly interviews and examines patients on his or her own.[c]	96.3	95.6	0.7	.758
3.12	Ask for the student's assessment and plan before giving your own formulation.[c]	95.1	100.0	−4.9	.009
3.13	Seek out the student to demonstrate physical findings on patients not seen by the student.[c]	92.6	83.8	8.8	.018
3.14	Ask questions to lead the student to his or her own diagnosis or treatment.[c]	92.6	91.2	1.4	.655
3.15	Regularly teach physical examination techniques.[c]	88.9	89.7	−0.8	.821
3.16	Watch the student do focused components of the physical examination (e.g., knee examination) to determine his or her skill level and learning needs.[c]	88.3	94.2	−5.9	.079
3.17	Create opportunities for the student to educate patients.[d]	85.9	78.7	7.2	.101
3.18	Help students identify uncertainty and formulate questions relating to patients.[d]	83.4	91.9	−8.5	.030
3.19	Create opportunities for the student to watch you manage difficult patient encounters.[d]	83.3	85.3	−2.0	.644
3.20	Create opportunities for the student to watch you communicate with patients.[d]	81.5	92.6	−11.1	.005
3.21	Give student time to organize his/her thoughts before they present their findings.[d]	78.5	77.4	1.1	.810

(Continued on next page)

TABLE 3
Thirty-two teaching behaviors valued by both students and preceptors, ranked within domains according to student responses (Continued)

No.	Behavior	% Respondents Valuing the Behavior			
		Students[a]	Preceptors[b]	Absolute Difference	*p*
Domain: Teaching Knowledge					
3.22	When a student incorrectly answers a question, don't leave the discussion there, but direct the student to the correct answer.[c]	95.1	94.8	0.3	.902
3.23	Take time during or immediately after each patient visit to ask if the student has questions or to make a teaching point.[c]	88.3	91.0	−2.7	.438
3.24	Use questions to help students improve their understanding of particular issues.[c]	87.7	94.0	−6.3	.064
3.25	Ask questions to probe the student's knowledge.[c]	76.1	89.0	−12.9	.004
Domain: Feedback					
3.26	Give the student an honest assessment of whether he or she falls short of any performance goal.[d]	95.7	93.4	2.3	.374
3.27	In feedback, do not stop at global criticisms. Be specific & directive, citing alternative ways of doing the pertinent skill.[d]	94.5	97.8	−3.3	.149
3.28	After telling the student of a skill, knowledge area, or attitude he or she needs to improve, help the student to improve.[d]	93.9	94.9	−1.0	.713
3.29	Follow negative criticism with action to help the student improve his or her performance.[d]	93.3	97.8	−4.5	.064
3.30	When students do something well, tell them they did it well.[d]	89.6	97.1	−7.5	.012
3.31	Give feedback during or after individual patient visits, not just during special sessions outside clinic hours.[d]	84.6	90.4	−5.8	.130
3.32	If a student does something wrong, tell him or her how to do it right. On the next occasion when the student does it correctly, complement him or her.[d]	88.3	95.6	−7.3	.023

[a] $N = 163$. [b] $N = 138$. [c] Identical or very similar items were identified as valued by students in a previous survey.[10] [d] Item was identified only during faculty focus groups. [e] Identical or very similar to items that were identified as not valued by students in a previous survey.[10]

disappoint the student if he or she views it as interference. With discussion beforehand, the student may understand that observation is a necessary basis for feedback and accept or even appreciate this occasional behavior.

In addition to discordant behaviors, our study identified a large number of specific behaviors ($N = 32$) that were valued by both medical students and their preceptors. Eight of the commonly valued behaviors were identified exclusively from focus groups of faculty preceptors. These 8 may not have been identified in student groups because of deficiencies in how the groups were conducted (e.g., not enough of them or inadequate methods), because students had not encountered them, or because students did not notice them. We believe the latter two explanations are more likely because student groups were conducted until no new behaviors emerged. Most of the 8 behaviors, furthermore, involve role modeling and educational design that students may not recognize as distinct teaching behaviors. The distinct contribution from preceptor focus groups indicates the importance of seeking input from both learners and teachers for research on practical aspects of education in ambulatory care locations.

Since 2000 when our earlier survey was published, two additional reports have examined medical students' perceptions of effective teaching behaviors.[14,15] Investigators at the University of Pittsburgh asked students to rate preceptors on 14 teaching behaviors. Multivariate analysis was used to identify 7 behaviors that were independently related to a rating of overall teaching effectiveness.[15] These 7 included behaviors (e.g., "preceptor treated student with trust and respect," "ethical medicine was practiced") that are broadly defined and difficult to compare to the more specifically defined behaviors that were the focus of our research. However, our findings complement one of the 7 broadly defined behaviors, helping the student learn clinical skills, by identifying specific teaching behaviors that preceptors could employ to succeed within this broader area (Table 3, Behaviors 3.9–3.21). Investigators at the Medical College of Wisconsin asked students to answer questions regarding individual patient encounters during an internal medicine clerkship.[14] In multiple logistic regression analysis, two teaching behaviors were related to higher overall rating of the teaching encounter: receiving high-quality feedback and

TABLE 4
Eighteen teaching behaviors not valued by both students and preceptors, ranked within domains according to student responses

No.	Behavior	% Respondents Valuing the Behavior			
		Students[a]	Preceptors[b]	Absolute Difference	*p*
Domain: Orientation to the Rotation					
4.1	Orient the student to the medical record.[c]	66.3	72.8	−6.5	.223
4.2	Introduce the student to everyone who works in the practice.[c]	58.9	55.9	3.0	.600
4.3	Early in the rotation, ask the student what experiences he or she hopes to have.[c]	58.0	69.9	−11.9	.035
Domain: Creating a Favorable Learning Environment					
	None				
Domain: Overseeing the Student's Experience					
4.4	Create in advance a daily list of patients who will be seen by the student—do not just select patients from your list.[d]	18.4	17.4	1.0	.827
Domain: Orchestrating the Student–Patient Interaction					
4.5	Hold preliminary discussions about diagnosis & treatment away from the patient.[e]	66.9	64.0	2.9	.599
4.6	Obtain consent from the patient for the student's participation.[c]	45.7	55.3	−9.6	.101
4.7	Before each patient encounter, give the student a specific time limit for completing the history and physical examination.[c]	29.0	43.0	−14.0	.012
Domain: Teaching Clinical Skills					
4.8	Delegate responsibility to the student for ascertaining and interpreting test results.[e]	82.5	67.6	14.9	.003
4.9	Leave the student alone with the patient until he or she has completed his or her evaluation.[d]	74.1	74.2	−0.1	.974
4.10	Facilitate the student's sense of being the caregiver.[e]	74.1	83.0	−8.9	.065
4.11	Have the student observe you caring for patients so that you can role model what you want them to do in your practice.[c]	69.8	84.7	−14.9	.002
4.12	Delegate responsibility to the student for telephone calls to patients (i.e., to check on treatment outcome or convey test results).[d]	48.1	43.3	4.8	.407
Domain: Teaching Knowledge					
4.13	Put students in the teaching role. Give them assignments to educate both of you.[c]	73.0	80.7	−7.7	.117
4.14	Choose reading assignments that are relevant: that influence patient care or educate other caregivers.[e]	72.2	83.0	−10.8	.028
4.15	Reserve time outside the clinic sessions to discuss patients with the student.[c]	60.2	66.4	−6.2	.274
4.16	Question students about their medical knowledge in front of patients.[c]	7.4	4.4	3.0	.286
Domain: Feedback					
4.17	Set a regular time to meet with the student to review patients and give feedback.[d]	74.1	73.0	1.1	.833
4.18	Watch the student do the visit/consultation closure.[c]	68.9	80.1	−11.2	.028

Note: Omitted from this table are the eight behaviors valued differently by student and preceptors which are listed in Table 2.
[a] $N = 163$. [b] $N = 138$. [c] Item was identified only during faculty focus groups. [d] Identical or very similar to items that were identified as not valued by students in a previous survey.[10] [e] Identical or very similar items were identified as valued by students in a previous survey.[10]

being asked to propose a plan. Both were also identified in our research.

Our research measured the value that students and preceptors assign to specific teaching behaviors; measurement of the actual effectiveness of the behaviors would require a different methodology. Other limitations of our research include its geographic focus in selected schools of the northeastern United States. We cannot be certain that our findings apply to schools in other regions, although this seems likely. Our survey did not include all possible specific teaching behaviors; we wanted to keep it short, left out some behaviors from our prior survey, and did not include input from professional educators. It is possible, therefore, that discordance or agreement may exist for other teaching behaviors used in ambulatory care environments. Finally, calculation of disagreement for the average value students and preceptors place on teaching behaviors may underestimate the burden of disagreement between individuals.

During ambulatory care clerkships students acquire professional competency under the supervision of preceptors who provide access to patients, graduated responsibility, and clinical instruction. The matrix for this experience is effective communication and collaboration between student and preceptor. The findings from this research indicate that this communication and collaboration should now involve matters of educational format and teaching behavior. The findings also describe a core set of teaching behaviors that should probably be part of every preceptor's routine.

REFERENCES

1. Kalet A, Schwartz MD, Capponi LJ, et al. Ambulatory versus inpatient rotations in teaching third-year students internal medicine. *Journal of General Internal Medicine* 1998;13:327–330.
2. Pangaro L, Gibson K, Russell W, et al. A prospective randomized trial of a six-week ambulatory medicine rotation. *Academic Medicine* 1995;70:537–41.
3. Harris IB, Watson K, Howe R. Development and evaluation of a required ambulatory medicine clerkship. *Academic Medicine* 1991;9:511–2.
4. Kenny NP, Mann KV, MacLeod H. Role modeling in physician's professional formation: Reconsidering an essential but untapped education strategy. *Academic Medicine* 2003;78:1203–10.
5. Lesky LG, Wilkerson L. Using "standardized students" to teach a learner-centered approach to ambulatory precepting. *Academic Medicine* 1994;69:955–7.
6. Skeff KM, Stratos GA, Berman J, et al. Improving clinical teaching. Evaluation of a national dissemination program. *Archives of Internal Medicine* 1992;152:1156–61.
7. Wilkerson L, Sarkin RT. Arrows in the quiver: Evaluation of a workshop on ambulatory teaching. *Academic Medicine* 1998;73:S67–9.
8. Green ML, Gross C, Kernan WN, et al. Integrating teaching skills and clinical content in a faculty development workshop. *Journal of General Internal Medicine* 2003;18:468–74.
9. Salerno SM, O'Malley PG, Pangaro L, et al. Faculty development seminars based on the one-minute preceptor improve feedback in the ambulatory setting. *Journal of General Internal Medicine* 2002;17:779–87.
10. Kernan WN, Lee MY, Stone SL, et al. Effective teaching for preceptors of ambulatory care: A survey of medical students. *American Journal of Medicine* 2000;108:499–502.
11. Morgan DL. *Focus groups as qualitative research* (2nd ed.). Thousand Oaks, CA: Sage, 1997.
12. Rogers HD, Carline JD, Paauw DS. Examination room presentations in general internal medicine clinic: Patients' and students' perceptions. *Academic Medicine* 2003;78:945–49.
13. Buchel T, Edwards F. Characteristics of effective clinical teachers. *Family Practice* 2005;37:30–5.
14. Torre DM, Sebastian JL, Simpson DE. Learning activities and high-quality teaching: Perceptions of third-year IM clerkship students. *Academic Medicine* 2003;78:812–4.
15. Elnicki DM, Kolarik R, Bardella I. Third-year medical students' perceptions of effective teaching behaviors in a multidisciplinary ambulatory clerkship. *Academic Medicine* 2003;78:815–9.

PREVENTING CHRONIC DISEASE
PUBLIC HEALTH RESEARCH, PRACTICE, AND POLICY
VOLUME 5: NO. 2 APRIL 2008

ORIGINAL RESEARCH

A New Brief Measure of Oral Quality of Life

Nancy R. Kressin, PhD, Judith A. Jones, DDS, MPH, Michelle B. Orner, MPH, Avron Spiro III, PhD

Suggested citation for this article: Kressin NR, Jones JA, Orner MB, Spiro A III. A new brief measure of oral quality of life. Prev Chronic Dis 2008;5(2). http://www.cdc.gov/pcd/issues/2008/apr/06_0147.htm. Accessed [*date*].

PEER REVIEWED

Abstract

Introduction

We developed a brief measure of the impact of oral conditions on individual functioning and well-being, known as oral quality of life.

Methods

Among older male veterans (N = 827) and community dental patients (N = 113), we administered surveys consisting of extant oral quality of life items, using clinical dental data from the veteran samples. We assigned each oral quality of life item to a theoretical dimension, conducted an iterative series of multitrait scaling analyses to examine the item-fit with the dimensions, reduced the number of items, and examined the psychometric characteristics of new scales and their association with clinical indices.

Results

We developed two brief oral quality of life scales, one consisting of 12 items and the other of 6, the latter a subset of the former. Each demonstrated sound psychometric properties and was sensitive to clinical indices.

Conclusion

The two brief oral quality of life scales can be used to assess the population-based impact of oral conditions as well as outcomes of dental care.

Introduction

The individual and public health impact of dental disease is increasingly recognized (1). However, dentistry has traditionally used specific clinical indices (e.g., number of teeth, periodontal attachment loss) to assess the impact of dental conditions. The limitations of using such clinical assessments of oral health status to understand the impact of oral disease are now clear (2): oral conditions affect the full scope of health status, including patients' functioning and well-being (e.g., oral quality of life [OQOL]) (3). Numerous patient-based measures of OQOL (4-8) have been developed, along with several clinician-based assessments (9-11). These measures vary in length (and, thus, ease of use in large-scale population-based surveys), in their sensitivity to clinical indices (or changes therein), and in their theoretical anchoring. Few studies have simultaneously examined the performance of items from more than one instrument (12).

Our goal was to produce a brief measure of oral health-related quality of life that was theoretically anchored and psychometrically and clinically valid, using best-performing items from existing instruments, to provide a public health tool for assessing the individual and population impacts of oral health conditions.

Methods

Samples

We studied two groups of older male veterans from the Veterans Health Study and the Dental Longitudinal Study. The medical and oral health status of these men covered a broad range of conditions. We conducted a brief clinical oral exam (<15 minutes) and administered an oral health-related quality of life questionnaire to men in each

study. In addition, to test our new brief instrument, we collected new questionnaire data on a sample of community dental patients at the time of a dental office visit.

1. The Veterans Health Study (VHS), begun in 1993, is a large-scale observational study (N = 2425) of veterans (mean age at the time of the study, 62 years), who are representative of independent, community-dwelling veterans who use Department of Veterans Affairs (VA) ambulatory care; as such, they had a variety of chronic medical conditions (13,14). A subset of 538 men completed the OQOL measures as part of an auxiliary study (15).

2. The Dental Longitudinal Study (DLS) is the dental component consisting of 1231 participants from the VA Normative Aging Study (NAS), a closed-panel longitudinal study of aging begun in 1968 among 2280 community-dwelling male veterans (mean age at the time of the study, 67 years) (16,17). The DLS is focused on oral health in aging men (18). Most are veterans but are *not* VA patients and are generally representative of the adult male population in the greater Boston metropolitan area, although they may be healthier, because all subjects were required to be medically healthy for entry into the NAS. DLS participants receive clinical dental exams triennially and were asked to complete an additional questionnaire concerning their oral quality of life; 289 participated.

3. Community dental patients who were visiting one of 16 community dental offices for prophylaxis, endodontic care, or placement of a removable prosthesis were provided our OQOL survey; 113 patients, both male and female, participated.

Data collection from each sample was approved by the local institution's institutional review board, and all participants gave written informed consent.

Theoretical framework

Our work was guided by a broad conceptualization of health and quality of life widely used in the general and oral health-related quality of life literature, which includes five broad domains (19-22):

- *Survival* is equated to mortality or longevity of the tooth or orofacial structures.
- *Impairments and disease* (or illness) include symptoms and indications of discomfort or pain.
- *Functional states* include three domains: a) *physical* functioning (e.g., activity restrictions, difficulties eating, chewing, or speaking); b) *social* functioning (e.g., the ability to perform social roles such as speaking, smiling, eating in public, being able to meet work and family obligations); and c) *psychological* functioning (e.g., patient satisfaction with the aesthetics of their dentition; comfort with interpersonal relations; worry, concern, embarrassment about, or lack of confidence caused by problems with teeth or gums).
- Oral health *perceptions* include one's global assessment of, and satisfaction with, oral health status and aesthetics, including need for treatment.
- Finally, *opportunity and resilience* reflect disadvantages incurred as a result of oral health and the impact of the disadvantages on one's ability to function in social and work roles and to have good nutrition through a satisfactory ability to eat and chew.

We adapted this model to fit an oral/dental framework and then examined how well the adapted model fit our data on oral health and quality of life. We hypothesized a framework with four primary dimensions: 1) physical function, 2) psychosocial functioning (with three subdimensions: role function, distress, and worry), 3) impairment or disease, and 4) perceptions.

Measures

The survey that we administered to our two veteran study populations included three extant OQOL measures (the Geriatric Oral Health Assessment Instrument, the Oral Health-Related Quality of Life [OHQOL] measure, and the Oral Health Impact Profile) with a total of 64 OQOL items, concurrent with a clinical oral exam. The community dental patient sample received a survey with an abbreviated selection of OQOL items but no clinical dental exam.

The Geriatric Oral Health Assessment Instrument (GOHAI) (5) comprises 12 items reflecting 3 domains of impact of oral disease: physical function, psychosocial function, and pain or discomfort.

The OHQOL measure is a brief global assessment of the impact of oral conditions on an individual's functioning and well-being (4). The three OHQOL items assess the extent to which problems with teeth or gums influence an

individual's daily activities, social interactions, or avoidance of conversations.

Slade and Spencer (23) developed the Oral Health Impact Profile (OHIP), an empirically grounded 49-item instrument based on a conceptual framework of oral disease and its functional and psychosocial consequences. The OHIP contains seven subscales: 1) functional limitation, 2) pain, 3) psychological discomfort, 4) physical disability, 5) psychological disability, 6) social disability, and 7) disadvantage.

Clinician-assessed oral health status

We collected clinical data in the VHS and DLS. To assess periodontal treatment need, we used the Community Periodontal Index of Treatment Need (CPITN) (24-26), which is based on measures taken from 10 teeth from the 6 sextants of the mouth, yielding an index score (ordinal scale of 0–4). This index was developed by the World Health Organization as an efficient measure for use in epidemiologic studies of periodontal status and treatment needs. *Coronal caries* and restorations were scored as in protocols of the National Institute of Dental Research for its National Survey of Oral Health in Adults (27), whereas *root caries* measures used an index developed by Hayes and Katz (28). This latter methodology, which has been used in two large epidemiologic studies, is efficient because it requires assessment of root caries and restorations on only eight tooth surfaces instead of every tooth in the mouth (29,30).

Procedure and data analysis

First, three of the authors (NK, JJ, AS) independently categorized each item from the three OQOL instruments into one of the theoretical domains described above. Any differences in domain assignments were resolved by consensus. Next, using existing data from the two veteran samples (N = 827), we conducted a series of psychometric analyses and examined the fit of the items to the hypothesized domains, using numerous iterations of multitrait scaling analysis (30,31), which is built on the logic of the multitrait-multimethod approach (32). Multitrait scaling analyses examine item-level characteristics (e.g., amount of missing data, frequency distribution, mean, standard deviation), the relationship of each item to other items in the scale that the item is hypothesized to measure, as well as the item's relationship to other scales. This analytic method provides information about scale distribution characteristics (e.g., mean, standard deviation, range, percentage of respondents scoring at the floor and ceiling) as well as the reliability of the scale scores and correlations among hypothesized scales. Compared with exploratory factor analysis, another commonly used approach to scale development, multitrait scaling analyses take a more confirmatory approach, evaluating the appropriateness of a priori groupings of items, allowing the investigator to specify and analyze conceptually meaningful groups of items. Item internal consistency (the extent to which the item is related to the concept being measured) is considered acceptable if an item correlates 0.40 or more with its hypothesized scale, after correction for item-scale overlap (30).

Item discriminant validity (the extent to which the item measures what it is supposed to measure) is considered acceptable if the correlation between the item and its hypothesized scale is significantly higher than the item's correlation with all other scales (32); we used the significance level of two standard errors (95% confidence interval) for this criterion. For internal consistency reliability (the extent to which items within a scale share common variance), we considered a Cronbach α of .70 to be acceptable (33).

Multiple approaches may be used to produce short-form measures of health-related quality of life (15), including item impact studies, factor analytic approaches (described above), and stepwise regression analysis. We were unable to adopt an item impact approach because we did not have item impact data for two of the three measures we used, and our choice of the multitrait scaling analysis was largely driven by our desire to confirm and refine our hypothesized conceptual schema.

Results

On the basis of the initial multitrait scaling analyses, we identified three items (GOHAI3, GOHAI5, GOHAI8) that correlated poorly with all of the domains we originally hypothesized, so we eliminated these items. We created a separate denture subscale, recognizing that denture functioning represents a conceptual dimension separate from that of natural teeth; this also further improved scaling properties. Results indicated that the psychosocial and opportunity items covered four dimensions: 1) distress,

2) self-consciousness and worry, 3) role function, and 4) opportunity. Because most of the items in the latter construct loaded more strongly on other scales and because of skepticism about the usefulness of opportunity as an oral health construct, we deleted these three items (OHIP29, OHIP45, OHIP47). Thus, we were left with three remaining psychosocial constructs: distress, self-consciousness and worry, and role function. Additional analyses found that some items had poor loadings on the hypothesized dimensions. Accordingly, we moved the perception items (OHIP44, OHIP3, GOHAI7) from the perception dimension into the worry dimension of psychosocial items, where they had higher loadings.

We then examined the impairment items, using exploratory factor analysis, because of concerns about the multidimensionality of this domain. Indeed, we found four subdimensions: 1) mouth pain, 2) flavor, digestion, and breath, 3) tooth pain, and 4) denture discomfort. On the basis of these results, we retained all of these items but further altered our conceptual model to include five dimensions: 1) physical function, 2) impairment and disease, and three dimensions of psychosocial function: 3) role function, 4) distress, and 5) worry (Appendix 1).

Using the remaining items, we standardized the item scores so that the mean of each variable was 50 and the standard deviation, 10. We scored the scales by taking the mean of all the items, after reversing the response categories where necessary so that higher scores indicated poorer oral quality of life. Thus, we created five scales to correspond with the above dimensions, a separate scale of the three denture-related items, and a summary scale comprising all items.

To develop a shorter version of the measure, we used data from the two veteran samples analyzed together to conduct forward stepwise regressions on each scale. This process allowed us to determine which items explained the most variance in each scale score. We selected items that explained either 80% of the variance or the first five items, whichever was greater. This resulted in five scales, each with five items. All of the scales had excellent internal consistency reliability, ranging from .78 (impairment) to .92 (distress), with the other scales also having excellent reliability (Table 1).

Next, we examined the correlations of each scale with clinical indices (Table 1). The strongest correlation observed was between physical function and number of teeth (r = −0.38). Coronal caries was moderately associated with worry (r = 0.23) and impairment (r = 0.18), whereas periodontal status was moderately associated with physical function (r = 0.21) and worry (r = 0.21). Root caries had the smallest correlations overall with OQOL.

We also examined mean scores on each of the quality of life dimensions by scores on the CPITN and found that individuals with greater treatment need had significantly worse OQOL (Table 2).

We then examined the proportion of variance explained in each oral quality of life dimension among different subgroups based on number of teeth (not shown). The impairment, physical function, worry, and role function scales explained the least variance among patients with no teeth and the most among patients having 1 to 10 teeth. The patterns observed for the distress scale were different: the most variance was explained among those with either no teeth or 1 to 10 teeth, with the least (but still 97% of the variance explained) among those with 11 to 24 teeth.

Next, we administered these five scales and the three denture-specific items to the sample of community dental patients. Using multitrait analysis, we sought to reduce the number of items further by eliminating items contributing least to each scale's internal consistency reliability and retaining items that conceptually best represented the spirit of the subscale. We eliminated items whose deletion least affected the internal consistency reliability of the scales (Cronbach α), and at the same time, sought to retain the items that we considered, from a conceptual standpoint, best represented the spirit of the subscale. We did this on two levels. First, we developed one 12-item measure (Appendix 2) that includes 3-item subscales for each of 3 scales in the psychosocial dimension (distress, worry, and role) and single items assessing dimensions entitled *physical, denture,* and *pain* (Cronbach α of the 12 items = .90). We also developed a second, briefer 6-item measure that includes single items assessing each dimension (distress, worry, role, physical, denture, pain) (Cronbach α for the scale = .80).

We then took these two brief measures, refined on the community dental patient sample, and returned to our original data set of 827 veterans to examine the association of the two brief scales with clinical indices. Both summary scales were significantly correlated overall with

number of teeth (r = −0.35 and −0.23, for the 6- and 12-item scales, respectively), coronal decay (r = 0.09 and 0.14), periodontal status (r = 0.19 and 0.20), and root caries (r = 0.14 and 0.12) (Table 3). Most items were significantly correlated with number of teeth, coronal decay, and periodontal status, but fewer were significantly correlated with root caries. Most items were associated with periodontal treatment need (Table 4).

Discussion

We sought to develop a brief oral quality of life measure that is theoretically anchored, psychometrically sound, and clinically responsive from items comprising three existing OQOL indices, and that can be used by public health researchers, practitioners, and policy makers to assess the impact of oral conditions on people's functioning and well-being. We conducted extensive psychometric analyses, reducing the original 64-item pool to a set of 25 items comprising 5 theoretically derived scales that demonstrate sound psychometric properties and associations in the expected directions with clinical indices. We further reduced the number of items to two brief scales (one with 12 items and one with 6 items) that maintained strong psychometric characteristics. Both scales were sensitive to differences in clinical dental status, supporting their validity.

Taken together, the findings indicate that these new oral quality of life measures are sensitive to clinical indicators of oral health status, suggesting their usefulness in monitoring population health, for making prevalence estimates, for monitoring secular trends in population changes, and for studying the effects of public health interventions designed to prevent or reduce the effects of oral disease. The associations we detected between the oral quality of life measure and clinical indices are similar to those of other published findings (12,34).

The intermediate results of examining the properties of the five 5-item scales showed that each scale accounted for a suitably high proportion of the variance. Thus, we conclude that the conceptual domains are well represented by the items in each scale. Importantly, our results indicate that both short-form scales are also sensitive to differences in clinical status and would be feasible to use in the clinical setting as an outcomes measure or in the general population to assess the impact of differing clinical status. The observed differences in internal consistency reliability suggest that the 6-item measure is appropriate for comparing groups of people, whereas the 12-item measure would be appropriate for assessing outcomes among individuals.

These results were limited by our partial reliance on cohorts of veterans who were older and all men, and who are thus not representative of the population as a whole. However, this disadvantage was offset by the availability of a rich clinical data set on these cohorts. Furthermore, the absence of detailed sociodemographic data on the community dental sample limited our ability to examine associations with these factors.

What value does this new brief OQOL instrument add to the literature, especially given that there are other OQOL instruments of similar length (e.g., OHIP14, GOHAI)? Other authors have compared the performance of various OQOL measures in their entirety (e.g., the OHIP14 and the OIDP [36,37] or the GOHAI and OHIP14 [38]), but to our knowledge, none have evaluated the relative performance of individual items from multiple OQOL measures. Our new measure has undergone extensive psychometric analysis and evaluation regarding its sensitivity to clinical indices and, although others have conducted similar types of analyses (36,38) using less detailed clinical data, the availability of extensive clinical dental data on our cohorts provides added confidence in this measure's sensitivity to differences in oral health status. Slade (38) examined the relative performance of the 49 original OHIP items and developed a 14-item short form of the OHIP. Two of the items in his short-form measure also are present in ours (finding it difficult to relax with oral problems and being totally unable to function because of oral problems), but several other OHIP items in our measure did not survive his item-reduction process, and 12 items in his short-form measure did not survive our item-reduction efforts.

Efforts are under way to use these new measures among populations of dental patients (39), including evaluations of responsiveness to change in clinical status (40). Future research would also benefit from a comparison of the performance of this new brief measure to that of other OQOL measures of similar length in multiple settings, including the community, and in private dental offices or as a treatment outcomes measure for use by dental insurers.

Acknowledgments

This research was supported by a grant from the Delta Dental Plans Association, the National Institutes of Health (NIH)/National Institute of Dental and Craniofacial Research grant number U54 DE14264-02, and by grants from the U.S. Department of Veterans Affairs Health Services Research and Development Service (IIR 93.025 Judith A. Jones, Principal Investigator; HFP 91-012, Pantel Vokonas, Principal Investigator; RCS 02-066-1, Nancy R. Kressin, Principal Investigator), and supported in part by NIH grants U54 DE014264, K24 DE000419 (R. I. Garcia, Principal Investigator), and K24 DE018211 (Judith A. Jones, Principal Investigator).

The views in this article are those of the authors and do not necessarily represent the views of the Department of Veterans Affairs or of the Delta Dental Plans Association.

Author Information

Corresponding Author: Nancy R. Kressin, PhD, Research Career Scientist, Center for Health Quality, Outcomes and Economic Research, 200 Springs Road, Building 70 (152), Bedford, MA 01730. Telephone: 781-687-2949; E-mail: nkressin@bu.edu.

Author Affiliations: Judith A. Jones, Boston University School of Dental Medicine; Center for Health Quality, Outcomes and Economic Research, Veterans Administration Medical Center (VAMC), Bedford, Massachusetts. Michelle B. Orner, Center for Health Quality, Outcomes and Economic Research, VAMC, Bedford, Massachusetts. Avron Spiro III, Center for Health Quality, Outcomes and Economic Research, VAMC, Bedford, Massachusetts; Normative Aging Study, Department of Veterans Affairs; and Boston Healthcare System, Boston University Goldman School of Dental Medicine, Boston University School of Public Health, Boston, Massachusetts.

References

1. National Institute of Dental and Craniofacial Research. Oral health in America: a report of the Surgeon General. Rockville (MD): U.S. Department of Health and Human Services; 2000.
2. Cohen L, Jago JD. Toward the formulation of socio-dental indicators. Int J Health Serv 1976;6(4):681-98.
3. Kressin NR. Symposium on self-reported assessments of oral health outcomes. Introduction. J Dent Educ 1996;60(6):485-7.
4. Kressin N, Spiro A III, Bosse R, Garcia R, Kazis L. Assessing oral health-related quality of life: findings from the Normative Aging Study. Med Care 1996;34(5):416-27.
5. Atchison KA, Dolan TA. Development of the Geriatric Oral Health Assessment Index. J Dent Educ 1990;54(11):680-7.
6. Slade GD, Spencer AJ. Development and evaluation of the Oral Health Impact Profile. Community Dent Health 1994;11(1):3-11.
7. Leao A, Sheiham A. The development of a socio-dental measure of dental impacts on daily living. Community Dent Health 1996;13(1):22-6.
8. Slade GD, editor. Measuring oral health and quality of life. Chapel Hill (NC): University of North Carolina, Dental Ecology; 1997.
9. Nikias MK, Sollecito WA, Fink R. An oral health index based on ranking of oral status profiles by panels of dental professionals. J Pub Health Dent 1979;39(1):16-26.
10. Marcus M, Koch AL. A proposed index of oral health status: a practical application. J Am Dent Assoc 1983;107:729-33.
11. Marcus M, Koch AL, Gershen JA. Construction of a population index of adult oral health status derived from dentists' preferences. J Public Health Dent 1983;43(4):284-94.
12. Jones J, Kressin N, Spiro A III, Miller D, Kazis L, Garcia R. Comparison of patient-based oral health outcome measures. Qual Life Res 2004;13(5):975-85.
13. Kazis LE, Miller DR, Clark J, Skinner K, Lee A, Rogers W, et al. Health related quality of life in VA patients: results from the Veterans Health Study. Arch Intern Med 1998;158:626-32.
14. Miller DR, Skinner KM, Kazis LE. Study design and sampling in the Veterans Health Study. J Ambul Care Manage 2004;27(2):166-79.
15. Jones J, Kressin N, Spiro A III, Randall C, Miller D, Hayes C, et al. Self-reported and clinical oral health in users of VA health care. J Gerontol A Biol Sci Med Sci 2001;56(1):M55-62.
16. Bell B, Rose CL, Damon A. The Veterans Administration longitudinal study of healthy aging. Gerontologist 1966;6(4):179-84.
17. Bosse R, Ekerdt DJ, Silbert JE. The Veterans Administration Normative Aging Study. In: Mednick

SA, Harway M, Finello KM, editors. Handbook of longitudinal research, Volume 2, teenage and adult cohorts. New York (NY): Praeger; 1984.
18. Kapur KK, Glass RL, Loftus ER, Alman JE, Feller RP. The Veterans Administration longitudinal study of oral health and disease: methodology and preliminary findings. Int J Aging Hum Dev 1972;2:125-37.
19. Patrick DL, Bergner M. Measurement of health status in the 1990s. Annu Rev Public Health 1990;11:165-83.
20. Patrick DL, Erickson P. Health status and health policy: quality of life in health care evaluation and resource allocation. New York (NY): Oxford University Press; 1993.
21. Gift HC, Atchison KA. Oral health, health, and health-related quality of life. Med Care 1995;33(11 Suppl): NS57-77.
22. Slade GD. The Oral Health Impact Profile. In: Slade GD, editor. Measuring oral health and quality of life. Chapel Hill (NC): University of North Carolina, Dental Ecology; 1997.
23. Locker D, Slade G. Association between clinical and subjective indicators of oral health status in an older adult population. Gerodontology 1994;11(2):108-14.
24. Ainamo J, Barmes D, Beagrie G, Cutress T, Martin J, Sardo-Infirri J. Development of the World Health Organization (WHO) Community Periodontal Index of Treatment Needs (CPITN). Int Dent J 1982;32(3):281-91.
25. Feldman RS, Douglass CW, Loftus ER, Kapur KK, Chauncey HH. Interexaminer agreement in the measurement of periodontal disease. J Periodontal Res 1982;17:80-9.
26. Ressler LL, Feldman RS. CPITN: Correlation and predictive value with treatment, plan and therapy. J Dent Res 1992;71:199.
27. The national survey of oral health in U.S. employed adults and seniors: 1985-1986. NIH Publication No. 87-2868. Bethesda (MD): National Institute of Dental Research; 1987.
28. Hayes C, Katz RV. An abbreviated examination for assessing root caries prevalence. J Dent Res 1992;71(Special Issue):250.
29. Hays RD, Hayashi R. Beyond internal consistency reliability: rationale and user's guide for multitrait analysis program on the microcomputer. Behav Res Methods Instrum Comput 1990;22(2):167-75.
30. Hays RD, Hayashi T, Carson S, Ware J. The Multitrait Analysis Program (MAP) user's guide. Santa Monica (CA): The Rand Corporation; 1988.
31. Campbell D, Fiske D. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychol Bull 1959;56(2):81-105.
32. Nunnally M, Bernstein I. Psychometric theory. 3rd ed. New York (NY): McGraw-Hill; 1994.
33. Jokovic A, Locker D, Guyatt G. Short forms of the Child Perceptions Questionnaire for 11–14-year-old children (CPQ_{11-14}): development and initial evaluation. Health Qual Life Outcomes 2006;4:4.
34. Robinson PG, Gibson B, Khan FA, Birnbaum W. Validity of two oral health related quality of life measures. Community Dent Oral Epidemiol 2003;31(2):90-9.
35. Duncan DB. *t* Tests and intervals for comparisons suggested by the data. Biometrics 1975; 31(2):339-59.
36. Robinson P, Gibson B, Khan F, Birnbaum W. A comparison of OHIP 14 and OIDP as interviews and questionnaires. Community Dent Health 2001;18(3):144-9.
37. Locker D, Matear D, Stephens M, Lawrence H, Payne B. Comparison of the GOHAI and OHIP-14 as measures of the oral health-related quality of life of the elderly. Community Dent Oral Epidemiol 2001;29(5):373-81.
38. Slade GD. Derivation and validation of a short-form oral health impact profile. Community Dent Oral Epidemiol 1997;25(4):284-90.
39. Cunha-Cruz J, Hujoel P, Kressin NR. Oral health-related quality of life of periodontal patients. J Periodontal Res 2007;42(2):169-76.
40. Wright W, Kressin NA, Jones JA. The use of outcome measures in dental education: the value of the OH-1 and DELTA. Paper presented at the 5th Joint Meeting of the Association of State and Territorial Dental Directors and the American Association of Public Health Dentistry; April 29, 2004; Los Angeles, California.

Tables

Table 1. Correlations of Scales (5 Items in Each) With Clinical Variables, Scale Internal Consistency Reliability, and Variance Explained by Each Scale, Among Participants in the Veterans Health Study and the Dental Longitudinal Study, 1993–1995 (N = 827)[a]

Scale	No. of Teeth (*P*)	Coronal Caries[b] (*P*)	Periodontal Status[c] (*P*)	Root Caries[d] (*P*)	Cronbach α	Variance Explained
Impairment	−0.02[e] (0.55)	0.18 (<.001)	0.19 (<.001)	0.13 (.004)	0.78	0.91
Physical	−0.38 (<.001)	0.10 (.011)	0.21 (<.001)	0.14 (.002)	0.81	0.94
Distress	−0.16 (<.001)	0.11 (.005)	0.14 (.002)	0.08[e] (.11)	0.92	0.97
Worry	−0.14 (<.001)	0.23 (<.001)	0.21 (<.001)	0.15 (.001)	0.84	0.93
Role	−0.19 (<.001)	0.08 (.02)	0.16 (<.001)	0.08[e] (.11)	0.86	0.96
Denture	−0.57 (<.001)	−0.09 (.01)	0.09[e] (.06)	0.09[e] (.05)	NA	NA
Summary scale	−0.21 (<.001)	0.16 (<.001)	0.21 (<.001)	0.13 (.01)	NA	NA

NA indicates not applicable.
[a] Pearson correlation coefficients were used to obtain means. Higher oral quality of life scores represent poorer quality of life.
[b] Coronal caries indicates coronal decayed surfaces at level 2 or greater.
[c] Periodontal status indicates per person mean Community Periodontal Index of Treatment Need (CPITN) score of available sextants.
[d] Root caries indicates mean percentage of exposed root surfaces with unfilled decay.
[e] Correlations are not statistically significant.

Table 2. Mean[a] Oral Quality of Life Scores by Varying Levels of Periodontal Disease Among Participants in the Veterans Health Study and the Dental Longitudinal Study, 1993–1995 (N = 827)

	CPITN Score[b]			
Scale	<1	1-1.9	2-2.9	≥3
All	48.0 (a)	47.4 (a)	51.4 (b)	53.9 (b)
Impairment	48.7 (a,b)	48.1 (a)	51.9 (b)	55.4 (c)
Physical	46.6 (a)	46.5 (a)	50.4 (b)	53.3 (b)
Distress	49.0 (a,b)	48.2 (a)	50.8 (a,b)	51.8 (b)
Worry	47.7 (a)	47.8 (a)	52.0 (b)	54.2 (b)
Role	49.2 (a,b)	47.9 (a)	50.9 (a,b)	52.7 (b)

CPITN indicates Community Periodontal Index of Treatment Need.
[a] Means were obtained from analysis of variance (ANOVA) testing and compared by using Duncan's multiple range test (35). Different letters indicate groups are significantly different from one another at $P < .05$; if the same letter is present, the groups are not different from one another. Thus, a mean labeled (a,b) is not significantly different from one labeled (b,c) because they both have a "b" beside them. Higher oral quality of life scores represent poorer quality of life.
[b] CPITN scores are as follows: <1, healthy periodontium; 1-1.9, gingival bleeding; 2-2.9, calculus; ≥3, moderate-deep periodontal pockets (need root planing).

Table 3. Correlations[a] of Items From the Two New Brief Scales and Overall Summary Scales With Clinical Variables Among Participants in the Veterans Health Study and the Dental Longitudinal Study, 1993–1995 (N = 827)

Scale	No. of Teeth (*P*)	Coronal Caries[b] (*P*)	Periodontal Status[c] (*P*)	Root Caries[d] (*P*)	Scale/Item No.
Impairment					
Past 3 months how much pain and distress[e]	−0.11 (.003)	0.11 (.004)	0.15 (<.001)	0.12 (.01)	OHQOL B31
Physical					
Have to avoid eating any food	−0.35 (<.001)	0.07[f] (.06)	0.17 (<.001)	0.14 (.004)	OHIP28
Distress					
Found it difficult to relax with oral problems	−0.13 (<.001)	0.09 (.01)	0.14 (.003)	0.08[f] (.08)	OHIP35
Feel depressed with oral problems	−0.16 (<.001)	0.10 (.009)	0.13 (.005)	0.06[f] (.18)	OHIP36
Being upset with oral problems	−0.12 (.001)	0.08 (.04)	0.12 (.01)	0.06[f] (.17)	OHIP34
Worry					
Uncomfortable about oral appearance	−0.12 (.001)	0.20 (<.001)	0.20 (<.001)	0.11 (.02)	OHIP22
Past 3 months feel nervous or self-conscious — teeth[e]	−0.15 (<.001)	0.14 (<.001)	0.18 (<.001)	0.09[f] (.08)	GOHAI10
Worried by dental problems	−0.06[f] (.13)	0.21 (<.001)	0.17 (<.001)	0.12 (.008)	OHIP19
Role					
Cannot get along with others	−0.13 (<.001)	0.08 (.04)	0.12 (.008)	0.05 (.27)	OHIP41
Avoid going out with oral problems	−0.14 (<.001)	0.11 (.004)	0.17 (<.001)	0.06 (.20)	OHIP39
Totally unable to function with oral problems	−0.11 (.003)	0.05 (.16)	0.11 (.02)	0.09 (.06)	OHIP48
Denture					
Have uncomfortable dentures	−0.53 (<.001)	−0.08 (.03)	0.05 (.28)	0.07 (.11)	OHIP18
Summary scale (6 items)	−0.35 (<.001)	0.09 (.009)	0.19 (<.001)	0.14 (.003)	NA
Summary scale (12 items)	−0.23 (<.001)	0.14 (<.001)	0.20 (<.001)	0.12 (.01)	NA

OHIP indicates Oral Health Impact Profile; GOHAI, Geriatric Oral Health Assessment Instrument; OHQOL, Oral Health Quality of Life; NA, not applicable.
[a] Pearson correlation coefficients were used to obtain means. Higher oral quality of life scores represent poorer quality of life.
[b] Coronal caries indicates coronal decayed surfaces at level 2 or greater.
[c] Periodontal status indicates per person mean Community Periodontal Index of Treatment Need (CPITN) score of available sextants.
[d] Root caries indicates mean percentage of exposed root surfaces with unfilled decay.
[e] Scores were reversed so that higher scores indicate poorer oral quality of life.
[f] Correlations are not statistically significant.

Table 4. Mean Oral Quality of Life Scores[a] by Varying Levels of Periodontal Treatment Need Among Participants in the Veterans Health Study and the Dental Longitudinal Study, 1993–1995 (N = 827)

Variables	CPITN Score[b]					Scale/Item No.
	≤1	1-1.9	2-2.9	3-3.9	*P* value	
Impairment						
Past 3 months how much pain and distress[c]	16.7 (b)	15.2 (b)	23.3 (a,b)	25.7 (a)	.003	OHQOLB31
Physical						
Have to avoid eating any food	13.9 (c)	15.9 (b,c)	23.4 (b)	34.1 (a)	<.001	OHIP28
Distress						
Found difficult to relax with oral problems	15.6 (a)	13.8 (a)	19.9 (a)	21.6 (a)	.04	OHIP35
Feel depressed with oral problems	17.0 (a)	13.8 (a)	21.9 (a)	21.0 (a)	.019	OHIP36
Being upset with oral problems	18.3 (a)	14.2 (a)	19.5 (a)	21.6 (a)	.124	OHIP34
Worry						
Uncomfortable about oral appearance	15.7 (b)	16.1 (b)	28.0 (a)	27.8 (a)	<.001	OHIP22
Past 3 months feel nervous or self-conscious — teeth[c]	8.9 (b)	10.2 (b)	20.2 (a)	23.2 (a)	.001	GOHAI10
Worried by dental problems	23.5 (b,c)	22.5 (c)	32.3 (a,b)	39.2 (a)	<.001	OHIP19
Role						
Cannot get along with others	8.3 (a)	8.4 (a)	11.2 (a)	14.8 (a)	.16	OHIP41
Avoid going out with oral problems	7.9 (a,b)	5.4 (b)	12.5 (a)	14.2 (a)	.003	OHIP39
Totally unable to function with oral problem	6.9 (a,b)	5.1 (b)	7.8 (a,b)	11.9 (a)	.07	OHIP48
Denture						
Have uncomfortable dentures	8.3 (a)	5.3 (a)	12.1 (a)	12.6 (a)	.02	OHIP18
6-item scale	12.4 (b)	10.8 (b)	18.4 (a)	21.8 (a)	<.001	NA
12-item scale	13.4 (b,c)	11.8 (c)	18.9 (a,b)	21.2 (a)	<.001	NA

CPITN indicates Community Periodontal Index of Treatment Needs; OHIP, Oral Health Impact Profile; GOHAI, Geriatric Oral Health Assessment Instrument; OHQOL, Oral Health Quality of Life; NA, not applicable.

[a] Means were obtained from analysis of variance (ANOVA) testing and compared by using Duncan's multiple range test (35). Different letters indicate groups are significantly different from one another at $P < .05$; if the same letter is present, the groups are not different from one another. Thus, a mean labeled (a,b) is not significantly different from one labeled (b,c) because they both have a "b" beside them. Higher oral quality of life scores represent poorer quality of life.

[b] CPITN scores are the following: <1, healthy periodontium; 1–1.9, gingival bleeding; 2–2.9, calculus; 3–3.9, moderate-deep periodontal pockets (need root planing).

[c] Scores were reversed so that higher scores indicate poorer oral quality of life.

Appendices

Appendix 1. Items Comprising Each 5-Item Oral Quality of Life Scale and the Denture Scale

Impairment
1. Have you had painful aching in your mouth? (OHIP9)
2. Have you felt that your sense of taste has worsened? (OHIP6)
3. Have you had sensitive teeth, for example, due to hot or cold foods or drinks? (OHIP12)
4. How much pain or distress have your teeth or gums caused you? (OHQOL B31)
5. Have you felt that your breath has been stale? (OHIP5)

Physical
1. Have you had to avoid eating some foods? (OHIP28)
2. Have you had trouble pronouncing any words? (OHIP2)
3. How often did you limit the kinds or amounts of food you eat because of problems with your teeth or dentures? (GOHAI1)
4. How often have problems with your teeth and gums affected your daily activities (such as work or hobbies)? (OHQOL1)
5. Have you found it uncomfortable to eat any foods? (OHIP15)

Distress
1. Have you found it difficult to relax? (OHIP35)
2. Have you felt depressed? (OHIP36)
3. Have you been a bit irritable with other people? (OHIP42)
4. Have you been upset? (OHIP34)
5. Have you been unable to enjoy other people's company as much? (OHIP46)

Worry
1. Have you felt uncomfortable about the appearance of your teeth, mouth, or dentures? (OHIP22)
2. How often did you feel nervous or self-conscious because of problems with your teeth, gums, or dentures? (GOHAI10)
3. How often have problems with your teeth and gums affected your social activities (such as with family, friends, coworkers)? (OHQOL2)
4. Have you avoided smiling? (OHIP31)
5. Have you been worried by dental problems? (OHIP19)

Role
1. Have you had trouble getting along with other people? (OHIP41)
2. Have you been unable to work to your full capacity? (OHIP49)
3. Have you avoided going out? (OHIP39)
4. Have you been totally unable to function? (OHIP48)
5. Have people misunderstood some of your words? (OHIP25)

Denture
1. Have you felt that your dentures have not been fitting properly? (OHIP17)
2. Have you had uncomfortable dentures? (OHIP18)
3. Have you been unable to eat with your dentures because of problems with them? (OHIP30)

OHIP indicates Oral Health Impact Profile; GOHAI, Geriatric Oral Health Assessment Instrument; OHQOL, Oral Health Quality of Life.

Appendix 2. Short Form 12-Item Oral Quality of Life Measure

During the past 3 months HOW OFTEN have you experienced the following difficulties because of problems with your teeth, mouth, or dentures? (Circle one answer)	**Never**	**Hardly Ever**	**Occasionally**	**Fairly Often**	**Very Often**
*1. Have you had to avoid eating some foods? (OHIP28)	0	1	2	3	4
*2. Have you found it difficult to relax? (OHIP35)	0	1	2	3	4
3. Have you felt depressed? (OHIP36)	0	1	2	3	4
4. Have you been upset? (OHIP34)	0	1	2	3	4
5. Have you felt uncomfortable about the appearance of your teeth, mouth, or dentures? (OHIP22)	0	1	2	3	4
6. Have you been worried by dental problems? (OHIP19)	0	1	2	3	4
7. Have you had trouble getting along with other people? (OHIP41)	0	1	2	3	4
*8. Have you avoided going out? (OHIP39)	0	1	2	3	4
9. Have you been totally unable to function? (OHIP48)	0	1	2	3	4

OHIP indicates Oral Health Impact Profile; GOHAI, Geriatric Oral Health Assessment Instrument; OHQOL, Oral Health Quality of Life.
* Indicates items included in 6-item measure.

(continued on next page)

Appendix 2. Short Form 12-Item Oral Quality of Life Measure (continued)

In the past 3 months, how often:	**Never**	**Sometimes**	**Always**
*10. Did you feel nervous or self-conscious because of problems with your teeth, gums, or dentures? (GOHAI10)	1	2	3

***11. During the past 3 months,** how much pain or distress have your teeth or gums caused you? (OHQOL B31)	**None at All**	**A Little Bit**	**Some**	**Quite a Bit**	**A Great Deal**
	1	2	3	4	5
If you have removable denture appliances, please answer the following question:					
During the past 3 months, how often have you had the following problems with your dentures?	**Never**	**Hardly Ever**	**Occasionally**	**Fairly Often**	**Very Often**
*12. Have you had uncomfortable dentures? (OHIP18)	0	1	3	4	5

OHIP indicates Oral Health Impact Profile; GOHAI, Geriatric Oral Health Assessment Instrument; OHQOL, Oral Health Quality of Life.
* Indicates items included in 6-item measure.

Perceptions of Race/Ethnicity-Based Discrimination: A Review of Measures and Evaluation of Their Usefulness for the Health Care Setting

Nancy R. Kressin, PhD
Kristal L. Raymond, MPH
Meredith Manze, MPH

Abstract: **Background**. To assess discrimination in health care, reliable, valid, and comprehensive measures of racism/discrimination are needed. **Objective**. To review literature on measures of perceived race/ethnicity-based discrimination and evaluate their characteristics and usefulness in assessing discrimination from health care providers. **Methods**. Literature review of measures of perceived race/ethnicity-based discrimination (1966–2007), using MEDLINE, PsycINFO, and Social Science Citation Index. **Results**. We identified 34 measures of racism/discrimination; 16 specifically assessed dynamics in the health care setting. Few measures were theoretically based; most assessed only general dimensions of racism and focused specifically on the experiences of African American patients. Acceptable psychometric properties were documented for about half of the instruments. **Conclusions**. Additional measures are needed for detailed assessments of perceived discrimination in the health care setting; they should be relevant for a wide variety of racial/ethnic groups, and they must assess how racism/discrimination affects health care decision making and treatments offered.

Key words: Discrimination, prejudice, delivery of health care, measurement.

Widespread racial/ethnic disparities in the quality of health care received, treatments offered, and health outcomes in the U.S. were documented in the Institute of Medicine report, *Unequal Treatment*,[1] but the reasons that racial/ethnic minority groups are likely to receive poorer quality care than Whites have not been fully explicated.[2–4] Hypothesized reasons for worse care include differentials in access to health care, actual health status, patient preferences, and provider bias or discrimination.[5,6] Yet, even after controlling for the former three factors, many studies have demonstrated that

NANCY KRESSIN *is affiliated with the Center for Health Quality, Outcomes and Economic Research (a VA Health Services Research and Development National Center for Excellence), Bedford VA Medical Center, Bedford, MA, and the Section of General Internal Medicine, Boston University School of Medicine, Boston, MA.* ***KRISTAL RAYMOND*** *was affiliated with the Department of Health Policy and Health Services Research, Boston University Goldman School of Dental Medicine, Boston, MA, at the time this research was conducted.* ***MEREDITH MANZE*** *is affiliated with Boston Medical Center, Boston, MA. Please address correspondence to Nancy R. Kressin, PhD, Section of General Internal Medicine, Crosstown Center, 2nd Floor, 801 Massachusetts Avenue, Boston, MA 02118; (617) 414-1367; nkressin@bu.edu.*

racial/ethnic minorities are less likely than non-Hispanic Whites to receive equivalent care across a broad spectrum of diseases, including cardiovascular disease, cancer, renal disease, asthma, mental illness, diabetes, and HIV/AIDS, suggesting that discrimination remains a plausible explanation.[1,3–5,7–9] Health care provider behaviors, attitudes, or treatment may vary according to patients' race/ethnicity. To assess such possible racism/discrimination and its effects, reliable and valid measures are needed.[1,3,6,7,10–17]

Three levels of racism have been described in the literature: personally-mediated, institutionalized, and internalized.[18] Personally-mediated racism, the focus of this paper, occurs via differential assumptions about the abilities, motives, and intentions of others according to their race/ethnicity (prejudice) and through differential actions toward others according to their race/ethnicity (discrimination). Racism, whether intentional or unintentional, may adversely affect the treatment of racial/ethnic minority patients in a variety of ways in the health care setting.[13] Health care provider bias can be as subtle as not giving certain patients the full range of treatment options, related to conscious or unconscious provider beliefs that certain patients are less willing to accept or adhere to certain therapies, or are incompetent, deviant, or not likeable.[3,6,9,14,18,19] Bias may also influence providers' clinical judgments[3] or their evaluations of a patient's personal or clinical characteristics.[14] African American physicians have argued that popular misconceptions, inaccuracies, and stereotypes of the psychology of African Americans could lead to misdiagnosis.[10]

Although it may be challenging for patients to discern whether they are experiencing discrimination, a growing body of literature deals with the methodology for measuring discriminatory behavior that is perceptible, and for inferring experiences of racism in a variety of settings, including community, school, workplace, and health care settings.[20,21] Prior reviews of this literature focused on measures that could be used specifically to assess the health effects of discrimination,[21] on comparing explicit versus generic measures of discrimination[22] or on the fewer measures of discrimination available at an earlier time.[20] However, the burgeoning literature on this topic suggests that an updated review of such measures is needed. The need for a new review is also bolstered by the fact that no previous review has evaluated measures' potential utility in the health care setting to assess accurately patients' experiences of racism or discrimination as they relate to the care received. Understanding the available measures is also important because until there is a reliable and valid method for assessing and comparing racism/discrimination across health care environments that can accurately assess the experiences of multiple population groups, it will not be possible to document the presence or degree of racism/discrimination, or to measure changes in its levels subsequent to interventions. Further, numerous new measures have appeared in the literature since the prior reviews; for example, the number of available measures has more than quadrupled since Utsey's 1998 review of six measures.[20] Thus, there is a need for an updated review and specific analysis of how well-developed or suitable such measures are for assessing experiences of racism in the health care environment, as well as the measures' appropriateness for use in different population groups within this setting.

The purpose of this review is to update and summarize the published literature on measures assessing individuals' perceptions of racism/discrimination in their environ-

ment (that is, an individual's appraisal of such dynamics in his/her environment or from people with whom s/he interacts) and to examine critically how these measures may apply when studying perceived racism in the health care setting. We propose that effective measures will assess whether racist/discriminatory events/actions occurred, the extent to which these bothered the individual experiencing them, and whether they affected the individual's interaction with his or her health care provider, including whether the experience affected the patient's view or acceptance of the provider's treatment recommendations,[12] or the provider's offer of care.

In addition, in order to understand the relevance of existing measures to a wide variety of population groups, we evaluate the racial/ethnic orientation of the measures (e.g., anchoring of an instrument to certain racial/ethnic groups' experiences that might be a function of belonging to a certain group). This compilation and critical review of existing measures and evaluation of their characteristics and gaps will be useful for guiding future researchers in their choice of measures to examine the provider contribution to racial/ethnic disparities in health care.

Methods

A comprehensive review of the medical and social science literature from 1966 to January 2007, using MEDLINE, PsycINFO, and the Social Science Citations Index, was conducted to identify relevant articles pertaining to perceived measures of racism or discrimination in the U.S. We sought to identify papers that simultaneously 1) addressed the dimensions of race and ethnicity (for which we used the search terms *race, ethnicity, Blacks, African Americans, Hispanic, Hispanic Americans, Latinos, Asian, Asian Americans, Native Americans, American Indians, ethnic groups, racial stocks, Caucasoid race,* or *Whites*); 2) addressed the issues of discrimination, racism, bias, unfair treatment, or prejudice (for which we used these exact terms); and 3) discussed the development or adaptation of a measure to assess the perceived experience of racism or discrimination (we excluded articles applying previously developed measures in order to capture unique measures). For the latter dimension we used the search terms *measure, measurement, discrimination measure*, and *measures of racial discrimination*. Using all databases, we limited our search to peer-reviewed articles in English that focused on experiences in the United States in human adults. When we conducted this search within Medline we identified 287 articles; after individually reviewing all entries to select only those directly relevant to this review, we constructed a list of 23 relevant papers. Within PsycINFO, the search identified 381 articles, which we pared down to 18 relevant papers. In the Social Sciences Citation Index, we identified 438 papers, retaining 16 for the review. Some of the discarded papers were not relevant or described studies using measures previously reported in other papers or did not provide sufficient information about the measure (e.g., a full list of the included items) to warrant inclusion in our review. In addition, a review of references within each article identified other relevant articles. Only studies that documented the initial presentation or refinement of structured written instruments to measure perceived racism, unfair treatment, or discrimination were included in the final analysis. After eliminating

overlap between the citation lists, 34 unique papers describing measures or adaptation of prior measures of racism or discrimination remained.

Analysis. We examined the measures of racism along five dimensions: 1) settings in which measures were developed/used, 2) theoretical frameworks, 3) content of measures, 4) psychometric qualities, and 5) populations studied (summarized in Table 1). For the measures that included items specific to the health care setting, we examined the number of health care-related items and the content of those items. We evaluated whether the measures examined the occurrence of specific racist/discriminatory events, the bothersomeness of the experience, or its effect on the individual's interaction with the provider, including assessments of the effects on the treatment received/recommended/accepted.

Results

As detailed in Table 1, 19 of the 34 articles included in this review describe an original instrument (one not previously introduced in the literature).[2,23–40] Fifteen describe adaptations, modifications, or further testing of previous instruments.[8,12,13,41–52] Instruments ranged in length from 1 item to 109 items.

Settings in which measures were developed/used. The majority of measures were oriented toward experiences of racism/discrimination in general or in the community setting.[2,8,23,26,28,30,31,36,38–42,44,45,48–51] Some measures were developed using student samples or were at least partially oriented toward educational settings[29,32,39,45,50,51] or the workplace.[24,25,33,34,39,50]

Sixteen measures directly assessed at least some dimensions of perceived racism in the health care setting.[8,12,13,23,26,35,36,38,39,42–44,46,48,50,52] One of these utilized the Perceptions of Racism Scale[36] while another used an adaptation of the Schedule of Racist Events,[43] while the others had few items specific to one's specific experience, and requested no information regarding treatment decisions. Two used similar general items.[13,35] Bird had the most detailed list of experiences in health care, however it focused on experiences someone had *ever* had in receiving health care rather than with a specific provider or in a specific care setting.[12] Numerous additional measures were designed to be employed in a variety of settings (such as occupational and educational) including a small number of items addressing experiences in the health care setting.[8,23,26,38,39,44,46,48,50]

Theoretical frameworks. We found 11 measures of perceived racism that were explicitly based on theoretical frameworks.[12,26,28,30–33,39,42,45,50]

The most common theoretical perspective taken by measure-developers was that of Lazarus and Folkman,[54] who stressed the importance of understanding individuals' perceptions of stressful experiences (this perspective was taken by eight of the instruments). Utsey,[20] for example, explicitly based his measures on this stress model. In this theoretical framework, the *interpretation* of (stressful) perceived racist events (major events and daily hassles) is more important than the objective events themselves,[54,55] because different individuals may appraise similar events differently, resulting in different effects on the individual (e.g., anger vs. self-doubt). Therefore, perception of racist events was the most important facet of the experience to assess.

Landrine posits that the *event and appraisal approach* is most appropriate for measur-

Table 1.

MEASURES OF RACIAL/ETHNIC DISCRIMINATION

Author	Study population	Study purpose and theoretical framework	Measure development and psychometric evaluation	Instrument description and examples of items
1. Amaro H., et al. (1987)	303 Hispanic women professionals via mailing of the National Network of Hispanic Women.	Experience of discrimination, job stress and peer support as related to mental health.	No psychometric evaluation described	109 closed-ended, self-administered questions on work, family, and health. Work characteristics: "Have you ever experienced discrimination?" Yes/No "How often do your peers undermine your performance?" (from (1) never to (5) always)
2. Auslander W.F., et al. (1997) (Healthcare)	158 (55 African American, 103 White) children with diabetes and their mothers or female guardians.	Whether perceived racism in the community was a predictor of mothers' satisfaction with medical care of their children with diabetes. Ecological framework to examine how family and community contexts may predict satisfaction with medical care.	Dressler's Perceptions of Racism Scale (PRS, see #6) was modified to measure perceptions of unfair treatment by workers in multiple settings including health care. The Cronbach's Alpha (CA)[a] coefficient was 0.78	6 questions, on a Likert-type scale Perceptions of unfair treatment on the basis of race by city officials, restaurant workers, health care providers, and school teachers.

(Continued on p. 702)

Table 1. ***(continued)***

Author	Study population	Study purpose and theoretical framework	Measure development and psychometric evaluation	Instrument description and examples of items
3. Bird S.T., et al. (2001) (Healthcare)	76 African American adults (59 women, 17 men) recruited from community centers and churches in northeast Ohio.	Perceived race-based and socio-economic status (SES)-based discrimination in interactions with healthcare providers. Also: Stigma consciousness questionnaire to capture African Americans' perceptions of being stigmatized when interacting with doctors and beliefs about how doctors view African Americans.	Adapted from a measure of "everyday discrimination" by Williams and colleagues (item #24). Stigma consciousness adapted from Pinel. Stigma consciousness questions 1-7: No psychometric evaluation Pinel: CA=0.78	(10 health care related items) Question: "When getting health care, have you ever had any of the following things happen to you because of your race or color?" 1–7 experiences were listed: 1) been treated with less courtesy than other people; 2) been treated with less respect than other people; 3) received poorer service than others; 4) had a doctor or nurse act as if he or she thinks you are not smart; 5) had doctor or nurse act as if he or she is afraid of you; 6) had a doctor or nurse act as if he or she is better than you; 7) felt like a doctor or nurse was not listening to what you were saying. Response format: yes/no Pinel: 7-point Likert-type 10-item questionnaire: ranging from 0=strongly disagree to 6=strongly agree, with 3=neither disagree or agree: 1) Stereotypes that doctors have about Blacks/AAs have not affected me personally; 2) I never worry that doctors will view my behaviors as stereotypically Black/AA; 3) I feel like doctors interpret all of my behaviors in terms of the fact that I am Black/AA; 4) Most doctors do not judge Blacks/AAs on the basis of their race; 5) My being Black/AA does not influence how doctors treat me; 6) I almost never think of the fact that I am Black/AA when I interact with doctors; 7) My being Black/AA does not influence how people of other ethnic groups act with me; 8) Most doctors have a lot more racist thoughts than they actually express; 9) I often think that doctors are unfairly accused of being racist; and 10) Most doctors have a problem viewing Blacks/AAs as equals.

(Continued on p. 703)

Table 1. *(continued)*

Author	Study population	Study purpose and theoretical framework	Measure development and psychometric evaluation	Instrument description and examples of items
4. Broman, C.L. (1996)	312 African American adults in Detroit.	Krieger 1990 study rephrased to refer only to discrimination in the past three years (see item #11).	No psychometric evaluation described	Study, stem questions were identical to Krieger 1990 measure and the situation varied slightly from Krieger (item #11).
5. Brondolo, et al. (2005) Perceived Ethnic Discrimination Questionnaire—Community Version	Study 1: 301 adults (174 African American, 82 Latino, 18 White, 7 Asian, 4 Native American, 9 mixed, 7 none reported) see in primary healthcare practices in NYC Study 2: 340 Black and Latino college students and adults	Study 1: To assess ethnic discrimination in any group, across varying levels of literacy, while assessing the life experiences of community-dwelling adults. Adaptation of Contrada's PEDQ. Based on Lazarus and Folkman's stress & coping framework Study 2: To develop brief version of Lifetime Exposure Scale	Study 1: Scale development: edited PEDQ (Contrada et al) to reduce vocabulary level and adapt items to experiences of community dwelling adults. Preliminary testing and interviews with small sample, then use in larger sample. Testing for construct validity Study 2: High internal consistency reliability remained; good convergent validity with other measures of racism; good construct validity	Study 1: 70 item questionnaire with 5 scales: Lifetime Exposure Discrimination, Discrimination in the Media, Discrimination Against Family Members, Discrimination in Different Settings (Five point response scale—never happened to happened very often) and Past Week Discrimination (four point scale: never in the past week to 3 or more times in the past week). Internal consistency reliability exceeded CA=.75 for all scales. **BECAUSE OF YOUR ETHNICITY/RACE …** How often … —Has someone said something disrespectful, either to your face or behind your back? —Have others ignored you or not paid attention to you? Study 2: 17 item brief version was created from the 34 item Lifetime Exposure scale from the PEDQ-CV. **BECAUSE OF YOUR ETHNICITY/RACE …** How often … —Have people been nice to you to your face, but said bad things about you behind your back? —Have others hinted that you must not be clean?

(Continued on p. 704)

Table 1. *(continued)*

Author	Study population	Study purpose and theoretical framework	Measure development and psychometric evaluation	Instrument description and examples of items
6. Clark, R. (2003) (Healthcare)	64 Black college students	Effects of perceived racism and social support on blood pressure reactivity.	Modified version of the 128-item Life Experiences and Stress Scale assessed perceived racism.	Perceived racism measurements: "Overall how much do you think that INTER-ETHNIC GROUP RACISM has had anything to do with any problem you have had related to your _____ in your lifetime?" (Nine domains included: employment, law enforcement and the legal system, money and finances, education, community, family and social relationships, emotional well-being, physical health and medical care, and public assistance.) Responses ranged from 0–4 (0=less than 25% of the time, 4=between 75% and 100% of the time). Perceived racism composite (PRC) reflected the sum of all responses.
7. Clark R., et al. (2006)	217 Black youth	Effects of perceived racism and coping responses on blood pressure.	Modified version of the 9-item Everyday Discrimination Scale to assess racism & 4-item assessment for coping responses	6-point Likert-type scale to assess racism: "In your day-to-day life how often have any of the following things happened to you BECAUSE OF YOUR RACE? [Responses ranged from 0 (almost everyday) to 6 (never).] Example: "you are called names or insulted" 4-item questionnaire used to assess coping response, 4-point Likert-type scale [Responses ranged from 1 (not at all) to 4 (a lot)]: "We are now interested in how you tend to respond to these types of experiences …" 1) "Tried to do something about it" 2) "Accepted if as a fact of life" 3) "Realized that you brought it on yourself" 4) "Talked to someone about how you were feeling"

(Continued on p. 705)

Table 1. *(continued)*

Author	Study population	Study purpose and theoretical framework	Measure development and psychometric evaluation	Instrument description and examples of items
8. Contrada, et al. (2001) Perceived Ethnic Discrimination Questionnaire	361 college students (208 White, 34 Black, 31 Hispanic, 60 Asian)	To develop a measure of discrimination that can be used in all ethnic groups. Based on Lazarus and Folkman's stress & coping framework.	To develop items, drew on conceptual analyses, journalistic descriptions, and qualitative analyses of subjects' descriptions of experiences of ethnic discrimination. Examined internal consistency reliability (all alphas>.70)	22 items assess seven forms of discrimination: verbal rejection, avoidance, exclusion, denial of equal treatment, disvaluing action, threat of aggression, aggression. Four scales with 7 point response (never – very often) developed: disvaluation, threat/aggression, verbal rejection and avoidance. Examples: How often have you been subjected to offensive ethnic comments aimed directly at you, spoken either in your presence or behind your back? How often has it been implied or suggested that because of your ethnicity you must be unintelligent?
9. Corrigan, P., et al. (2003) Modified Schedule of Racist Events (SRE) (Healthcare)	1,824 persons (285 African American, 33 Latino, 97 Native American, 15 Asian, 228 White) with serious mental illness recruited from community mental health centers. Multistate study.	Compared discrimination experienced by persons with mental illness to self-reports of discrimination due to other group characteristics such as race, gender and sexual orientation.	Measures were based on responses to the Discrimination Questionnaire, adapted from the Schedule of Racist Events (SRE). Modifications to SRE: changed focus to examine multiple sources beyond race and used global questions rather than behavior specific. No psychometric evaluation described	(one health care related item) 26 interview-based measures were administered: "Do you believe that you have been discriminated against, for instance because of your mental disability, race, gender, sexual orientation, economic circumstance or some other reason?" The next questions asked yes or no questions about the specific conditions of discrimination (one of which was race). The last questions asked about the type of situations the discrimination occurred (setting, including housing, employment, and education).

(Continued on p. 706)

Table 1. *(continued)*

Author	Study population	Study purpose and theoretical framework	Measure development and psychometric evaluation	Instrument description and examples of items
10. Dressler, W.W. (1990)	186 25–55 yr. old African Americans.	Effects of lifestyle incongruity (extent to which a high status style of life exceeds an individual's occupational class) on blood pressure. Effects of perceived stressors were evaluated.	No psychometric evaluation described	3 of 16 , yes or no, items from the Scale of Chronic Social Role Stressors: 1) "Feel you missed a promotion because you're Black?" 2) "Feel you're not given real job responsibilities because you're Black?" 3) "Feel you are paid less than a White person?"
11. Eccles J.S., et al. (2006)	1480 African American adolescents residing near Washington D.C.	Effects of racial discrimination on academic performance among African American 7th graders (longitudinal study with follow-up after completion of eighth grade)	Data from Maryland Adolescents Development in Context (MADIC) Study Parental perceptions of discrimination at work and in community by peers and teachers: Perceived discrimination by peers: CA: 0.86 Perceived discrimination by teachers: CA: 0.88	5-point scale for perceived discrimination 1) how often do you perceive poorer treatment in stores of restaurants because of your race? (ranging from 1=almost never to 5=almost daily) 2) compared to people of other races, how many opportunities for job advancement did you get at work (1=a lot fewer to 5=a lot more) Perceived discrimination by peers: examples include frequency they felt they got into fights, were not associated with, and were not picked for particular teams or activities because of their race. Perceived discrimination by teachers: examples include how often they felt their teachers called on them less, graded them more harshly, disciplined them more harshly, discouraged them from taking a class, and thought they were less smart because of their race. Questions also assessed adolescents' perception of future racial job and educational discrimination, racial group cultural-capital identity, achievement motivation, future endeavors, academic achievement, and their parents' perceptions of racial discrimination at work and in the community.

(Continued on p. 707)

Table 1. *(continued)*

Author	Study population	Study purpose and theoretical framework	Measure development and psychometric evaluation	Instrument description and examples of items
12. Green, N.L. (1995) Perceptions of Racism Scale (PoRS) (Healthcare)	African American Childbearing Women: Study A, N=109; Study B, N=136.	Extent racism was a factor in low-birth-weight (LBW) and preterm babies in African American women. Racism experienced: —Affectively (feelings of racism) —Behaviorally (racist actions) —Cognitively (racist thoughts)	Several items from the Business Week/ Harris Poll (Jackson and Collingsworth, 1988) were revised and included in the scale. CA: Study A=088 Study B=0.91	(9 health care related items) 20-item self-report inventory, of which 10 questions concern medical, 2 of life experiences of discrimination. 4-point Likert Type scale ranging from "strongly agree" to "strongly disagree": 1) AA women experience negative attitudes when they go to a White doctor's office; 2) Doctors treat AA and White women the same; 3) Racism is a problem in my life; 4) A pregnant White women is treated with more respect than a pregnant AA woman; 5) I am not affected by discrimination; 6) Sometimes if you are AA in a White doctor's office it's as if you don't belong there; 7) Racial discrimination in a doctor's office is common; 8) In most hospitals, AA women and White women get the same kind of care; 9) Doctors and nurses act the same way to White and AA pregnant women; 10) If an AA pregnant woman comes to a doctor's office, it's assumed that she is on welfare; 11) AAs have the same opportunities as Whites to live a middle class life; 12) Officials listen more to Whites than AAs; 13) If an AA woman and a White woman are applying for the same job they have the same chance of being hired; 14) There has been significant progress in ending racism in the 1980's; 15) A White woman has more opportunities than an AA woman; 16) AA women get pregnant to receive more welfare benefits; 17) AA woman can receive they want as equally as White women; 18) Judges are harder on AAs than Whites; 19) AA pregnant woman have fewer options for health care; 20) Officials listen more to AA's than Whites.

(Continued on p. 708)

Table 1. *(continued)*

Author	Study population	Study purpose and theoretical framework	Measure development and psychometric evaluation	Instrument description and examples of items
13. Harrell, J.P. (1994) Racism and Life Experiences Scales—Brief Version (RaLES-B)	139 African American/Latino Adolescents.	Assessed impact of perceived racism on behavior, psychological status, & health outcomes in minority populations. Developed in 1992 to assess the degree to which perceived racism influenced a group of African American and Latino men to participate in a substance abuse treatment program.	CA: Group=0.90 Self=0.88 Validity: Self was significant & positively correlated with immersion stage of racial identity (r=0.26) and adaptive functioning (r=0.18) group was significantly correlated with encounter (r=0.23) and internalization. (r=0.29) subscales of the RIAS-B.	32, 4-point Likert type scale Part 1— Perceived racism-self total How does it affect you? Part 2— Perceived racism-group total How does perceived racism affect your race?
14. Jackson, J.S., et al. (1996)	623 African Americans (national sample).	To investigate the consequences of racism for the physical and mental health of African Americans. Two measures of racism were used: 1) Perception of whites' intentions; and 2) A report of racial discrimination experiences.	No psychometric evaluation described	Perception of White's intentions response format: Select one of the 3 specified options: Whites want to keep Blacks down; Whites want to see Blacks get a better break; or, Whites just don't care one way or the other about Blacks. Discrimination response format: yes/no This measure asked Blacks whether they or their family member were treated badly because of their race in the past month.

(Continued on p. 709)

Table 1. *(continued)*

Author	Study population	Study purpose and theoretical framework	Measure development and psychometric evaluation	Instrument description and examples of items
15. James, et al. (1994) Workplace Prejudice/ Discrimination Inventory	Study 1: 90 'minority' volunteers from 4 different workplaces (>60% Mexican Americans) Study 2: 46 incumbents in one organization with multiple units (about half Mexican American, half Asian American)	Study 1: To develop a validated inventory to examine perceptions of racial prejudice and discrimination in the workplace, based on social identity theory. Study 2:	Study 1: 16 Items developed by examining literature, including individuals' global perceptions of prejudice/ discrimination at work and specific types of discrimination. Factor analyses revealed one primary factor with CA=.90 Study 2: Factor analysis led to one factor, CA=.93	7 item Likert response scale Examples: Where I work members of some racial/ethnic groups are treated better than members of other groups. At work I am treated poorly because of my racial/ethnic group.

(Continued on p. 710)

Table 1. *(continued)*

Author	Study population	Study purpose and theoretical framework	Measure development and psychometric evaluation	Instrument description and examples of items
16. Johnson, R.L., et al. (2004) (Healthcare)	6,299 White, African American, Hispanic and Asian adults.	Examined: 1) whether racial/ethnic differences exist in patients' perceptions of primary care provider and general health care system-related bias and cultural competence, and 2) whether these differences are explained by patient demographics, source of care, or patient-provider communication variables.	Data from the nationally representative Commonwealth Fund 2001 Health Care Quality Survey was used. No psychometric evaluation described	(6 health care related items) Physician Bias and Interpersonal Competence Measures (questionnaire Items): 1) Did the doctor treat you with a great deal of respect and dignity, a fair amount, not too much, or none at all? 2) I feel that my doctor understands my background and values. 3) I often feel as if my doctor looks down on me and the way I live my life. (For both 2 & 3: Strongly agree, somewhat agree, somewhat disagree, strongly disagree) Health System Bias and Cultural Competence Measures (questionnaire items): 1) Do you think there was ever a time when you would have gotten better medical care if you had belonged to a different race or ethnic group? 2) Thinking about all of the experiences you have had with health care visits in the last 2 years, have you ever felt that the doctor or medical staff you saw judged you unfairly or treated you with disrespect because of your race or ethnic background? 3) Thinking about all of the experiences you have had with health care visits in the last 2 years, have you ever felt that the doctor or medical staff you saw judged you unfairly or treated you with disrespect because of how well you speak English? (For questions 1–3, yes or no response)

(Continued on p. 711)

Table 1. *(continued)*

Author	Study population	Study purpose and theoretical framework	Measure development and psychometric evaluation	Instrument description and examples of items
17. Krieger, N. (1990) (Healthcare)	51 African American and 50 White women, aged 20–80, who resided in Alameda County, CA.	To determine the feasibility of asking questions pertaining to race- and gender-based treatment plus response to unfair treatment, and to assess their predictive value regarding self-reported high blood pressure.	No psychometric evaluation described	(one health care related item) Phone interview: Ever experienced discrimination response format: yes/no 1) If you feel you've been treated unfairly, how do you usually respond-do you: "Accept it as a fact of life?" "Try to do something about it?" 2) And if you've been treated unfairly, do you: "Talk to other people about it?" "Keep it to yourself?" 3) Have you ever experienced discrimination, been prevented from doing something, or been hassled or made to feel inferior in any of the following 5 situations because of your race or color? "At school," "getting a job," "at work," "at home," "getting medical care."
18. Krieger, et al. (2005) (Healthcare)	Cohort of working class adults, 159 African American, 249 Latino, and 208 White participants	To assess the validity and reliability of the "Experiences of Discrimination" measure.	Examined scale reliability through factor analysis, testing of internal consistency reliability (CA=.74), test-retest reliability (0.70), and found highest correlation with underlying discrimination construct compared to other self-report discrimination measures	(see items above)

(Continued on p. 712)

Table 1. *(continued)*

Author	Study population	Study purpose and theoretical framework	Measure development and psychometric evaluation	Instrument description and examples of items
19. Landrine, H., et al. (1996) Schedule of Racist Events (SRE) (Healthcare)	153 African American students, staff and faculty at a university.	Assessed frequency encountered racial discrimination, and degree to which a racist event was appraised as stressful, in past year and lifetime. Theoretical framework: Life Events (Dohrenwend, 1978) and Daily Hassles.	This measure was modeled after the PERI-LES and other similar scales that assess frequency of stressful events in people's lives. Recent discrimination (past year): CA=0.95, split-half reliability: 0.93 Lifetime discrimination: CA=0.95, split-half reliability: 0.91 Appraisal of stress: CA=0.92, split-half reliability: 0.92	(one health care related item) 18 item self-report inventory measures, on a 6-point Likert-type scale, the frequency African Americans have experienced specific racist events in a wide variety of settings, including one item regarding health care setting (e.g., racism experienced by institutions, neighbors, peers, and teachers). Each question was asked three times: once for the frequency of the racist event in the past year, once for the frequency of the event during one's entire lifetime, and once for the appraisal of the stressfulness of the events. Examples: "How many times have you been treated unfairly by institutions (schools, universities, law firms, the police, the courts, the Department of Social Services, the Unemployment Offices and others) because you are Black?" How many times have you been treated unfairly by people in helping jobs (doctors, nurses, psychiatrists, case workers, dentists … therapists …) Frequency: "never" to "almost all the time" Stress: "not at all" to "extremely" "How many times in the past year?" "How many times in your entire life?" "How stressful was this for you?"

(Continued on p. 713)

Table 1. *(continued)*

Author	Study population	Study purpose and theoretical framework	Measure development and psychometric evaluation	Instrument description and examples of items
20. Landrine H., et al. (2006) (Healthcare)	1569 adults (868 college students and 701 community adults; 780 Whites, 406 Latinos, 174 African American, 94 Asian Americans, 95 Other racial background)	To measure perceived ethnic discrimination in health research	General Ethnic Discrimination Scale modeled on the Schedule of Racist Events (SRE) that measures discrimination as a type of stress. Recent discrimination (past year): CA=0.94 Lifetime discrimination: CA=0.94 Appraised discrimination: CA=0.95 (split-half reliability for each scale: 0.91)	18-item measure of perceived ethnic discrimination, on a 6-point Likert-type scale , the frequency the sample has experienced specific racist events in a wide variety of settings, including one item regarding health care setting (e.g., racism experienced by institutions, neighbors, peers, and teachers). Each question was asked three times: once for the frequency of the racist event in the past year, once for the frequency of the event during one's entire lifetime, and once for the appraisal of the stressfulness of the events. Examples: "How many times have you been treated unfairly by institutions (schools, universities, law firms, the police, the courts, the Department of Social Services, the Unemployment Offices and others) because of your race/ethnic group?" How many times have you been treated unfairly by people in helping jobs (by doctors, nurses, psychiatrists, case workers, dentists … therapists …) because of your race/ethnic group? Frequency: "never" to "almost all the time" Stress: "not at all" to "extremely" "How often in the past year?" "How often in your entire life?" "How stressful was this for you?"

(Continued on p. 714)

Table 1. *(continued)*

Author	Study population	Study purpose and theoretical framework	Measure development and psychometric evaluation	Instrument description and examples of items
21. Lillie-Blanton, et al. (2000) (Healthcare)	Nationally representative sample of 3,884 Whites, African Americans and Latinos	To assess the public's attitudes about racial and ethnic differences in health care	Survey was designed and analyzed by the authors.	(three health care related items) Have you felt a doctor or health provider judged you unfairly or treated you with disrespect because of race or ethnic background?
22. Mays and Cochran (1997)	232 African American women and 73 African American men in college, university, in junior college, in Los Angeles, CA.	Frequency of discrimination as related to psychological distress.	Frequency of discrimination: —Based on race-ethnicity, gender or both: in general: personally experienced —As perpetrated by three sources (African American men, women, White men): against African American person of same gender as respondent; personally experienced —As perpetrated by other African Americans against African Americans lacking economic resources: in general; personally experienced No psychometric evaluation described	Frequency of discrimination response format: for each item, 7-point Likert-type scale, ranging from "never" to "fairly often." Degree of upset and relation to perpetrator, for each type of personally experienced discrimination, response format: 7-point Likert-type scale. Upset: ranging from "not at all" to "upset a great deal." Relationship to perpetrator: "mostly by those I know well" to "mostly by complete strangers"

(Continued on p. 715)

Table 1. *(continued)*

Author	Study population	Study purpose and theoretical framework	Measure development and psychometric evaluation	Instrument description and examples of items
23. McNeilly, M.D., et al. (1996) Perceived Racism Scale (PRS) (Healthcare)	165 African American College students, and 25 members from the community.	This measure of African Americans experience of racism provides an assessment of frequency of exposure to multiple types of racist events and an assessment of emotional and behavioral coping responses to racism across a variety of situations.	CA range: 0.88–0.96 Frequency of exposure=0.96 Emotional and behavioral coping responses=0.92 Test-retest reliability: range=0.05–0.78	(one health care related item) 51 items Measures frequency of exposure (past year; over one's life) to types of racist incidents, and includes dimensions of emotional and behavioral coping responses. These exposures and responses are measured with respect to their occurrence in three situational domains: on the job, in academic settings and in the public realm (overt and subtle), & measurement of exposure to racist statements which can occur across settings. One item regarding health care: "I have been denied hospitalization or medical care" **Dimension 1. Frequency of exposure to types of racist events** Types include: Individual and institutional, overt and covert, attitudinal and behavioral, temporal component assesses incidents occurring: —over the past year —over one's life **Dimension 2. Emotional responses to perceived racism** Emotions include feeling: Angry, frustrated, sad, powerless, hopeless, ashamed, strengthened. **Dimension 3. Behavioral coping responses to perceived racism** Coping behaviors include: Speaking up, accepting it, ignoring it, trying to change things, keeping it to myself, working harder to prove them wrong, praying, avoiding it, getting violent, forgetting it, other. **Domains—Each dimension is assessed across the following domains:** Employment domain, academic domain, public domain, racist statements

(Continued on p. 716)

Table 1. *(continued)*

Author	Study population	Study purpose and theoretical framework	Measure development and psychometric evaluation	Instrument description and examples of items
24. Peterson, et al. (2004)	1,979 full time medical school faculty. 82% White, 8% Black, 1% Mexican Americans, 1% Puerto Ricans, 0.3% Native Americans, 7% Asian or Pacific Islanders, 1% Hispanic Americans	To assess experiences of discrimination of minority faculty in academic medicine	Not described	Do you perceive any racial/ethnic biases or obstacles to the career success or satisfaction of faculty by race/ethnicity in your academic environment? (1=no, never – 5= yes, frequently) In your professional career, have you ever been left out of opportunities for professional advancement based on race/ethnicity? (1=no; 5=yes) In your professional career, have you personally encountered racial/ethnic discrimination (unfair or injurious distinction or treatment by a superior or colleague? (1=no; 2=yes) If yes to latter, respondents asked "How much of a problem has this been for you?," "Have you encountered racial/ethnic remarks?", "To what extent have these experiences had a negative effect on your confidence as a professional?", and "To what extent have these experiences negatively affected your career advancement?"
25. Ren, X.S., et al. (1999) (Healthcare)	The National Survey of Functional Health (NSFH) White=1525 African American=134 Hispanics=46 Asian or Others=42	Analyzed self-perceived unfairness (discrimination due to racial identity or to low SES) was linked to self-assesses health status.	Mail out survey adapted from Krieger, 1990 (item #12): added 2 more scenarios, and changed "discrimination" to "unfair treatment." CA coefficient of 0.83	1-Item: "Have you ever experienced unfair treatment, been prevented from doing something, or made to feel inferior because of race in 7 different situations: at school, getting a job, at work, getting medical care, getting housing, from the police or in the courts, and on the street or in a public setting?"

(Continued on p. 717)

Table 1. *(continued)*

Author	Study population	Study purpose and theoretical framework	Measure development and psychometric evaluation	Instrument description and examples of items
26. Ryan, A.M., et al. (2006) (Healthcare)	666 adults residing in New Hampshire (78 African American, 112 Black immigrants, 476 Latino immigrants)	Relationship between perceived discrimination and health (specifically blood pressure)	Data from the 2002–2003 New Hampshire Racial and Ethnic Approaches to Community Health 2010 Initiative (NH REACH)	(one health care related item) Questions adapted from the Reactions to Race module from the 2002 Behavior Risk Factor Surveillance System 1) How often do you feel discomfort or anger by the way others treat you in your everyday life because of your race? (constantly; once/day; once/week; once/month; once/year; never; other, specify) 2) Do you feel that racial discrimination diminishes your ability to achieve your goals fully? (yes/no) 3) Do you feel that you have been receiving less than the best health care because of your race? (yes, often; yes, some of the time; no, none of the time; don't know)
27. Salgado de Snyder, V.N. (1987)	Mexican immigrant women N=140, ages 17–49 (M=25.7 yrs) with a mean of 9.5 years of education, married for the first time in LA county.	Assessed acculturative stress, individual stressors and their relationship to levels of depressive symptomatology.	CA coefficient of 0.65	1 of 12 yes or no items on an acculturation scale, if the subject experienced stressor in the past 3 months. If the answer was yes, subjects were asked to further respond on a 4-point Likert type scale for degree of stressfulness (0—not very stressful, 4—very stressful).

(Continued on p. 718)

Table 1. *(continued)*

Author	Study population	Study purpose and theoretical framework	Measure development and psychometric evaluation	Instrument description and examples of items
28. Thompson, V.L. (1996)	200 African American adults.	To learn whether perceived racism produces symptoms of subjective distress noted in relationship to other stressful life events. Also, to find whether racial identification mediated the psychological impact of perceived experiences of racism. Stressful Life Events (Lazarus)	Racism: Inter-rater reliability= 89.6%	Racism Measure: Participants reported whether they had ever experienced and whether they had experienced racism in the last 6 months. Racism was described as an unfavorable, unfair, or insulting event or action that occurred due to their skin color or group membership. Three examples were provided if, and only if, a participant requested it (e.g. Loss of job due to race, refusal of housing due to race, or derogatory names that were racial in nature). Participants were asked to provide a written description of the most recent racial incident within 6 months. The raters were trained by the primary investigator to categorize racial incidents. Descriptions were categorized as minor, moderate and major.

(Continued on p. 719)

Table 1. *(continued)*

Author	Study population	Study purpose and theoretical framework	Measure development and psychometric evaluation	Instrument description and examples of items
29. Thompson, C.E., et al. (1990) Racism Reaction Scale (RRS)	European American (EA) college students (n=70); African American college students in Southern California (n=87); Total (N=157).	The study aim was to examine the experiences of African American college students attending predominantly EA universities. The measure for racism was designed to capture the individual level of suspicion about being singled out for differential treatment or being personally threatened.	The instrument used consisted of the CMI. CA of 0.86	The CMI questionnaire consisted of 2 subscales: —The Education and Training subscale was designed to measure the degree to which the respondent agrees that EAs are to be trusted by AAs within the context of education and training. —The Interpersonal subscale was designed to measure the extent to which respondents agree that AAs should mistrust EA's in interpersonal situations. Respondents on both subscales were instructed to respond to a 7-point Likert-type scale (ranging from 1=strongly disagree to 7=strongly agree): "Professors treat me differently than they do other students" "Professors don't expect me to perform as well as other students"
30. Trivedi, A.N., et al. (2006) (Healthcare)	54,968 adults (African American, White, Latino, Asian, AI)	Perceived discrimination in health care and use of preventive health services	Data from the 2001 California Health Interview Survey (CHIS)	Phone interview: 1) Thinking of your experiences with receiving health care in the past 12 months, have you felt you were discriminated against for any reason? 2) What do you think was the reason you were discriminated against? (10 response options included age, race, language, health and disability, weight, insurance, income, gender, medical beliefs, multiple reasons, and other).

(Continued on p. 720)

Table 1. *(continued)*

Author	Study population	Study purpose and theoretical framework	Measure development and psychometric evaluation	Instrument description and examples of items
31. Utsey, S.O. (1996) Index of Race Related Stress (IRRS)	African American adults with a wide range of demographic backgrounds: College students (CS) & community residents (CR) (pilot n=377); CS, CR and drug treatment clients (Study 1, n=302); CS&CR (Study 2, n=153).	Race-related stress operationalized as the occurrence and perceived magnitude of specific events of racism and discrimination potentially experienced in daily lives. —Cultural —Institutional —Individual —Collective Daily hassles (Lazarus and Folkman '84) integrated with Essed's ('90) concept of every day racism.	Initial items included on the IRRS was based on informal interviews with African Americans from diverse backgrounds, a review of the literature, and the personal life experiences of the researcher (an African American male). CA Study 1: Cultural racism=0.87; Institutional racism=0.85 Individual racism=0.84 Collective racism=0.79 CA Study 2: Cultural racism=0.89; Institutional racism=0.82 Individual racism=0.84 Collective racism=0.74	46 item, 5-point Likert-type scale If an individual was the victim of a racist or discriminatory act, they were to indicate their reaction to the event on the basis of the following response choices: 0=this event has never happened to me, 1=this event happened but did not bother me, 2=this event happened and I was slightly upset, 3=this event happened and I was upset, 4=this event happened and I was extremely upset. Individuals responded only to the events that they experienced. "Whites or other non-Blacks have treated you as if you were unintelligent and needed things explained to you slowly or numerous times" Stress = general concept like emotion, but operationalized according to the perceived magnitude of the interactions with the environment that taxed or exceeded the person's resources.

(Continued on p. 721)

Table 1. *(continued)*

Author	Study population	Study purpose and theoretical framework	Measure development and psychometric evaluation	Instrument description and examples of items
32. Utsey—brief version (IRRS-B) (1999)	College students (25 European Americans for comparison group and 239 African Americans)	To develop a shorter version of the IRRS.	CA's: Cultural racism=0.78; Institutional racism=0.69 Individual racism=0.78 Confirmatory factor analyses supported a three factor model; the brief measure had higher factor loadings, more robust fit indices, and adequate CA.	22 item subset from the original items, as described above.
33. Vines, A.I., et al. (2001) Telephone-Administered Perceived Racism Scale (TPRS)	476 employed African American women: randomly selected members of an urban HMO.	Perceptions of and responses to racism among working African American women.	Some of the items were used from the Perceived Racism Scale (see item # 17) The emotional and behavioral responses were related to on the job and public experiences, and subscales for active and passive coping responses were utilized.	Experiences of racism (as a group and at the individual level), emotional responses, behavioral responses and past experiences of racism: Components of TPRS: Personal: Respondent discriminated when needing medical care Group: Blacks are paid less Passive emotions: Hopeless and powerless

(Continued on p. 722)

Table 1. *(continued)*

Author	Study population	Study purpose and theoretical framework	Measure development and psychometric evaluation	Instrument description and examples of items
33. Continued (Healthcare)			Alpha reliability values: 0.75 to 0.80 for the active and passive emotions subscale 0.59–0.69 for passive behaviors subscale >0.76 for both active behaviors subscales 0.82, 0.90, 0.88 & 0.82 for past experiences, concern for child(ren), experiences of racism (personal), and experiences of racism (group)	Active emotions: Angry, frustrated, sad Passive behaviors: Does not speak up or try to change things External active behavior: Working harder to prove them wrong Internal active behavior: Praying Concern for child(ren): Being punished more harshly than others in school Past experiences of racism: Frequency of past experience. Stress of past experience.
34. Williams D.R., et al (1997)	520 Whites and 586 African Americans from the Detroit Area Study.	Racial differences in SES, acute and chronic indicators of perceived discrimination, and general measures of stress account for differences in self-reported measures of physical and mental health. Essed (Everyday events).	Two measures of race-related stress: major discrimination and everyday discrimination. Both were framed in the context of "unfair" treatment rather than racism. CA=0.88 for Everyday Discrimination	Discrimination, as measure of major experiences, is a count of 3 items: 1) Do you think you have ever been unfairly fired or denied promotion? 2) For unfair reasons, do you think that you have ever not been hired for a job? 3) Do you think you have ever been unfairly stopped, searched, questioned, physically threatened or abused by police? Everyday discrimination (chronic, routine, and relatively minor experiences): Sums nine items on everyday situations. E.g. being treated with less courtesy than others; they think you are dishonest; they are afraid of you.

[a]Cronbach's Alpha = CA

ing racist events.[26] Appraisal refers to the assessment of the racist event and the psychoemotional resources available for dealing with them.[55] According to this author, ways of dealing with racism may influence perceived racism or its effects during the health care encounter. One ecologic framework, found in the work of Auslander, highlighted the importance of health behaviors and beliefs and ways in which individual behaviors and cognition about medical care are influenced by family and community.[42] Therefore, questions pertaining to family and community were included in the instrument in order to decipher perceived racism experienced within this broader context.

Others have assessed race-related stressful life events, explicitly building on Essed's notion of everyday discrimination—the idea that specific incidents of racial bias can affect one's well-being in a manner different from the effects of major experiences of unfair treatment.[12,31] The remaining instruments were presented without reference to a theoretical or conceptual base.

Content of measures. *Targets of race-based discrimination*. All of the measures of perceived racism recorded in Table 1 assess the individuals' experiences of situations and environments, as this was part of our selection criteria. We refer to this as *individual-as-target* measures; they asked questions about racism directed toward the individual, based on his or her race/ethnic group membership. In addition, one measure also asked about racism experienced by family members.[2]

Time span. Almost all of the 34 measures asked about *any* prior experiences of perceived racism from the target's perspective.[2,8,12,13,23–50] In addition, several used a second reference period for measurement; one used a two-year reference period,[13] three used a one-year reference period,[26,38,39] one used a six-month reference period,[28] one used a three-month reference period[32] and two used a one-month reference period.[2,51] Vines used the following time period for reference: before 20 years of age and during one's 20s.[8]

Responses to racism measured. A person who experiences discrimination responds with emotional or behavioral coping mechanisms, and interprets this experience according to individual belief systems.[38] In measuring experiences of discrimination, it is important to understand someone's response to such experiences (e.g., passive, active, behavioral, and emotional), whether expressed or suppressed, in addition to simply whether someone is exposed to racism, in order to evaluate its impact on the provision of health care and health outcomes. We found that 11 of the 34 publications reviewed asked about the type of responses to racism,[8,26–28,30,38,41,46,48,49] above and beyond specific experiences the individual had had. None, however, inquired about the impact of racism/discrimination on treatment decisions or care received in the health care setting.

Question(s) framed in context of unfair treatment based on other characteristics. Racism can either be measured by asking respondents to indicate specific experiences of racism (e.g., unfair treatment by doctors) or to respond to less focused questions about unfair treatment in general. Seven publications mentioned unfair treatment in the instruments as a reference to racism.[23,24,26,39,46,48,51] In order to compare unfair treatment based on race/ethnicity to unfair treatment because of gender or other factors, respondents to three measures were further asked about unfair experiences based on gender, socioeconomic position or social class, sexual preference, and religion.[23,46,52] Almost all of the measures specifically anchored the questions about discriminatory

treatment in the respondent's race/ethnicity rather than gender, sexual orientation, or another trait, with two exceptions.[24,52]

Psychometric evaluation. Over half (19 of 34) of the publications described some form of psychometric evaluation of the measures of perceived racism.[8,12,24,26–33,36,38–40,42,44,47,51] Krieger conducted the most detailed psychometric analyses of any of the measures, showing that the *Experiences of Discrimination* measure had high internal consistency reliability, good test-retest reliability, and correlated with other self-report discrimination measures.[46] Most of the psychometrically evaluated measures had Cronbach's Alpha scores of at least 0.70 in all aspects of the instrument, and thus were shown to have high internal consistency reliability, with the exception of three measures.[8,27,38]

Race/ethnicity of populations studied. The vast majority of studies included African American respondents.[2,8,12,13,23,25,26,28–32,34–52] Fourteen studies included Hispanic respondents.[13,24,27,32–34,39,40,43–46,48,52] Nine included Asian respondents.[13,32–34,39,43–45,52] In eight studies, the category *Latino* was specified,[35,39,40,43,45,46,48,52] and in two studies Mexican immigrants were included.[27,33] Native Americans were included as a race/ethnic category for four studies.[34,43,45,52]

Measures that pertain to the health care setting. We focused additional analytic attention on the 16 measures that included items for the specific purpose of examining individuals' experiences in the health care setting.[8,12,13,23,26,35,36,38,39,42–44,46,48,50,52] We evaluated the number of items focused on health care experiences, examined the content of the items, and reviewed whether they asked the degree to which the experience was bothersome. Sixteen original measures included at least one item about experiences in obtaining health care, but none included more than 10 items. In general, the content of the items focused on racist attitudes of doctors and other health care workers, and general discriminatory behaviors such as being treated unfairly, being treated with less respect, or being provided poorer service than other patients. McNeilly's measure included an item about being denied hospitalization or medical care because of race/ethnicity and Ryan's measure included an item about receiving less than the best health care because of race.[38,48] These were the only specific questions about the process of medical care in any of the measures. Seven of these measures assessed the bothersomeness of racist/discriminatory treatment.[26,30,37–39,47,48]

Discussion

Our evaluation of the 34 measures identified in the medical or psychological literature that assess individual perceptions of experiences of racism or discrimination in their environment revealed than only 16 such measures were specifically developed for or relevant to experiences in the health care setting. Of these measures, half included only a single item.[23,26,38,39,44,46,48,50] The remaining measures were limited by their sole focus on the experiences of African Americans without inclusion of experiences of other minority groups[12,36,42,43,52] or by their use of questions about only general experiences (e.g., being treated with less courtesy, disrespect, or poorer service).[8,13,35] We view these as significant limitations of the available measures. Without attention to how patients feel about their experiences of racism, or how bothersome the experiences of racism/discrimination were, or how such experiences specifically affect the medical

care provided, it is difficult to assess the impact of racism/discrimination on patients' care. Thus, the existing measures are useful to characterize the *general* experiences of patients of color, particularly African Americans, in obtaining medical care, similar to patient satisfaction measures' gestalt ratings of care, but the existing measures cannot inform the field regarding the *specific* impact of racism/discrimination on treatment recommendations or health care decisions.

We recommend that researchers refine and extend existing measures or develop new measures that specifically assess racism in the health care setting and the ways in which it affects medical care, since the ultimate impact of racism/discrimination in this setting differs from that of, for example, experiences in the general community.

Only 8 of the 34 publications reported asking respondents about the kinds of responses, whether active or passive, they had to racism.[8,23,30,38,46–49] We view this as a limitation in the majority of the existing instruments. Without an understanding of individuals' responses to racial discrimination, research cannot evaluate how troubling or significant the experiences were, nor elucidate how such experiences ultimately affected the provision of health care.

According to some researchers, measures should avoid global questions, due to the likelihood of underestimating exposure to experiences of racism.[21] Examples include the single-item global measure of racism[10,31] such as *ever* experiencing racism and questions using the term *unfair treatment*.[39] These questions may be too vague to apply to the race-based experience of the individual.[23,27,46] Furthermore, the generic experience of unfair treatment can be generalized to all people who have such experiences. We believe it is vital to anchor questions about unfair treatment to the issue of race/ethnicity specific to the health care encounter in order to provide specific information for policy implementation and interventions in the health care setting.

In addition to asking about individuals' experiences, it may be helpful to ask questions about family and/or group experiences, as some measures did.[38,50] Individuals may deny racism to avoid feeling that they do not have control over situations, and some prefer not to recall such memories.[56–59] A few studies indicate that some minority group members tend to minimize racism and attribute their failure to themselves (internalized racism).[58,60] Consequently, some experiences of racism may be underreported. Since minimizing recall of racism may be psychologically beneficial to the individual,[60] future instruments should be designed to be sensitive to the tendency of individuals to deny reports of racism, without undermining it.[40]

Past research indicates that *everyday discrimination*, often referred to as *chronic exposure* to racism, is a better predictor of health status than major (*acute*) experiences.[12,31,61] Furthermore, chronic perceived racism may affect health care utilization and subsequent health outcomes.[12] Thus, we recommend that measures include assessments of chronic racism and racism in society in general, because including them may provide important information about potential confounders and effect modifiers of the true racism exposure-health outcome relationship in the health care setting.[41]

Current measures of experiences of racism are subject to several methodological pitfalls. For instance, in order to measure accumulated exposure to racism, respondents are generally asked to indicate whether they have experienced racism within a certain time span. Yet, timing of experiences (e.g., *ever* or *past month*) may introduce recall

bias since people tend to forget experiences over time.[30,41] It may be more beneficial to include a one-year reference period when asking about racism, according to the literature on stressful life events.[2]

Comparison of measures was challenging because some focused on the frequency while others focused on the severity of racism.[31] People sometimes forget about the intensity of an experience over time,[41,62] and the nature of racism in the U.S. has changed over time as well, becoming less overt.[18,29,30,41,62,63] We recommend asking questions about both subtle and overt experiences of racism, within a specified time frame.

The absence of a clearly elucidated theoretical framework for many of the measures we reviewed is a matter of concern. To be able to measure the existence and extent of racism, it is necessary to have a theory of how racism might occur and what its effects might be. Without such a theory to guide measure development, analysts may conduct studies using invalid measures that do not have interpretable results. Thus, we recommend that future measures be anchored in theory, such as Lazarus and Folkman's view of stress, as described above.

Using the same measure across diverse groups of people can be conceptually and psychometrically problematic. As a result, issues and concerns addressed in the design of an instrument to assess perceived racism may not be relevant to the population being studied, because measures were not developed with this population in mind. Inclusion of information about subgroups within racial/ethnic minority populations may allow for better analysis of cultural influences on perceived experiences of racism.[10] Although nearly half of the measures reviewed in our collection of publications were psychometrically evaluated, most focused on the experiences of African American populations, and fewer focused on those of other racial/ethnic groups. As the U.S. population becomes increasingly diverse, due in part to the rapid growth of the Hispanic population, it will become imperative to assess accurately the experiences of various groups. We advocate continued development of new measures or enhancement of existing measures to ensure that psychometrically and conceptually valid instruments are applicable to more diverse population groups.

This review was limited in several ways. We focused on perceptions of racism and discrimination instead of more objective assessments (although, to our knowledge, few objective tools are available). Second, patient perceptions may not be the best tool with which to document structural or institutional discrimination, which are often invisible to individual participants. Thus, the fact that racial and ethnic minorities often receive care in poorer quality facilities undoubtedly leads to worse outcomes, but it would be hard for an individual to perceive such differences. Finally, individuals from racial/ethnic minority backgrounds may perceive poor treatment in the health care setting as being a function of race, which may not be the case. Despite these limitations, the data from this review document the state of the art in measuring racism/discrimination in the health care setting, and provide important direction for future measurement development and refinement.

Research on the contribution of provider behavior to disparities in medical care is in its infancy, and there have been few studies specifically designed to test the effect of provider and health care personnel behavior on these disparities.[3] Further study is needed to validate the hypothesis that provider behavior during encounters is indepen-

dently influenced by patient race/ethnicity.[3] Furthermore, most of the prior measures of perceived racism have focused on the experiences of African Americans, with less emphasis on other population groups. There is an obvious gap in the literature to address the potential contribution of provider behavior to disparities in medical care for American Indians/Alaskan Natives, Asians, and Hispanic populations. Therefore, to meet the challenge of explicitly measuring other racial/ethnic minority populations' experiences, new methods and approaches for measuring perceived racism in health care settings are needed. The information provided by such studies may help to design interventions intended to ameliorate the provider contribution to disparities in care for diverse racial/ethnic populations.

Acknowledgments

This work was supported by NIH/National Institute of Dental and Craniofacial Research grants U54 DE14264-02, and K24 DE00419, NIH/NHLBI grant R01 HL072814. Ms. Raymond received a Research Supplement to Promote Diversity in Health-Related Research in support of her work (U54 DE014264-04S2). Dr. Kressin is supported by a Research Career Scientist award (RCS 02-066-1) from the Health Services Research and Development Service, Department of Veterans Affairs.

The views expressed in this article are those of the authors and do not necessarily represent the views of the Department of Veterans Affairs.

Notes

1. Institute of Medicine of the National Academies. Unequal treatment: confronting racial and ethnic disparities in health care. Washington, DC: National Academy Press, 2002.
2. Jackson JS, Brown TN, Williams DR, et al. Racism and the physical and mental health status of African Americans: a thirteen year national panel study. Ethn Dis. 1996 Winter–Spring;6(1–2):132–47.
3. Van Ryn M. Research on the provider contribution to race/ethnicity disparities in medical care. Med Care. 2002 Jan;40(1 Suppl):I140–51.
4. Balsa AI, McGuire TG. Statistical discrimination in health care. J Health Econ. 2001 Nov;20(6):881–907.
5. Kressin N, Petersen L. Racial differences in the use of invasive cardiovascular procedures: review of the literature and prescription for future research. Ann Intern Med. 2001 Sep 4;135(5):352–66.
6. Kressin NR, Chang BH, Whittle J, et al. Racial differences in cardiac catheterization as a function of patients' beliefs. Am J Public Health. 2004 Dec;94(12):2091–7.
7. 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 Feb;93(2):248–55.
8. Vines A, McNeilly MD, Stevens J, et al. Development and reliability of a Telephone-Administered Perceived Racism Scale (TPRS): a tool for epidemiological use. Ethn Dis. 2001 Spring–Summer;11(2):251–62.
9. Snowden LR. Bias in mental health assessment and intervention: theory and evidence. Am J Public Health. 2003 Feb;93(2):239–43.

10. Williams DR, Williams-Morris R. Racism and mental health: the African American experience. Ethn Health. 2000 Aug–Nov;5(3–4):243–68.
11. South-Paul JE. Racism in the examination room: myths, realities, and consequences. Fam Med. 2001 Jun;33(6):473–5.
12. Bird ST, Bogart LM. Perceived race-based and socioeconomic status (SES)-based discrimination in interactions with health care providers. Ethn Dis. 2001 Autumn; 11(3):554–63.
13. Johnson RL, Saha S, Arbelaez JJ, et al. Racial and ethnic differences in patient perceptions of bias and cultural competence in health care. J Gen Intern Med. 2004 Feb;19(2):101–10.
14. Van Ryn M, Burke J. The effect of patient race and socio-economic status on physicians' perceptions of patients. Soc Sci Med. 2000 Mar;50(6):813–28.
15. Walters KL. Reconceptualizing native women's health: an "indigenist" stress-coping model. Am J Public Health. 2002 Apr;92(4):520–4.
16. American Public Health Association. Eliminating Health Disparities: Communities Moving from Statistics to Solutions—Toolkit.
17. Williams DR, Rucker TD. Understanding and addressing racial disparities in health care. Health Care Financ Rev. 2000 Summer;21(4):75–90.
18. Jones CP. Invited commentary: "race," racism, and the practice of epidemiology. Am J Epidemiol. 2001 Aug 15;154(4):299–304; discussion 305–6.
19. U.S. Commission on Civil Rights, Office of the General Counsel. Broken promises: evaluating the Native American health care system. Washington, DC: U.S. Commission on Civil Rights, 2004; 25–33.
20. Utsey SO. Assessing the stressful effects of racism: a review of instrumentation. J Black Psychology. 1998;24(3):269–88.
21. Krieger N. Embodying inequality: a review of concepts, measures, and methods for studying health consequences of discrimination. Int J Health Serv. 1999;29(2): 295–352.
22. Brown TN. Measuring self-perceived racial and ethnic discrimination in social surveys. Sociological Spectrum. 2001;21:377–92.
23. Krieger N. Racial and gender discrimination: risk factors for high blood pressure? Soc Sci Med. 1990;30(12):1273–81.
24. Amaro H, Russo NF, Johnson J. Family and work predictors of psychological well-being among Hispanic women professionals. Psychology of Women Quart. 1987;11: 505–21.
25. Dressler W. Lifestyle, stress, and blood pressure in a southern black community. Psychosom Med. 1990 Mar–Apr;52(2):182–98.
26. Landrine H, Klonoff EA. The schedule of racist events: a measure of racial discrimination and a study of its negative physical and mental health consequences. J Black Psychology. 1996;22(2):144–68.
27. Salgado de Snyder VN. Factors associated with the acculturative stress and depressive symptomatology among married Mexican immigrant women. Psychology of Women Quart. 1987;11:475–88.
28. Thompson VL. Perceived experiences of racism as stressful life events. Community Ment Health J. 1996 Jun;32(3):223–33.
29. Thompson CE, Neville H, Weather PL, et al. Cultural mistrust and racism reaction among African American students. J College Stud Dev. 1990;31:162–8.
30. Utsey SO, Ponterotto JG. Development and validation of the index of race-related stress (IRRS). J Counseling Psychology. 1996;43(4):490–501.

31. Williams DR, Yu Y, Jackson JS. Racial differences in physical and mental health. J Health Psychology. 1997;2(3):335–51.
32. Contrada R, Ashmore R, Gary M, et al. Measures of ethnicity-related stress: psychometric properties, ethnic group differences, and associations with well-being. J App Soc Psychology. 2001;31:1775–820.
33. James K, Lovato C, Cropanzano R. Correlational and known-group comparison: validation of a workplace prejudice/discrimination inventory. J App Soc Psychology. 1994;24(17):1573–92.
34. Peterson NB, Friedman RH, Ash AS, et al. Faculty self-reported experience with racial and ethnic discrimination in academic medicine. J Gen Intern Med. 2004 Mar; 19(3):259–65.
35. Lillie-Blanton M, Brodie M, Rowland D, et al. Race, ethnicity, and the health care system: public perceptions and experiences. Med Care Res Rev. 2000;57(Supplement 1):218–35.
36. Green N. Development of the perceptions of racism scale. Image J Nurs Sch. 1995 Summer;27(2):141–6.
37. Mays V, Cochran S. Racial discrimination and health outcomes in African Americans. Presented at: National Center for Health Statistics 1997 Joint Meeting of the Public Health conference on Records and Statistics and the Data Users Conference, Washington (DC), July 1997.
38. McNeilly MD, Anderson NB, Armstead CA, et al. The perceived racism scale: a multidimensional assessment of the experience of white racism among African Americans. Ethn Dis. 1996 Winter–Spring;6(1–2):154–66.
39. Landrine H, Klonoff EA, Corral I, et al. Conceptualizing and measuring ethnic discrimination in health research. J Behav Med. 2006 Feb;29(1):79–94. Epub 2006 Feb 10.
40. Harrell JP. The racism and life experience scale. Unpublished manuscript. 1994 (Los Angeles: California School of Professional Psychology).
41. Broman CL. The health consequences of racial discrimination: a study of African Americans. Ethn Dis. 1996 Winter–Spring;6(1–2):148–53.
42. Auslander WF, Thompson SJ, Dreitzer D, et al. Mothers' satisfaction with medical care: perceptions of racism, family stress, and medical outcomes in children with diabetes. Health Soc Work. 1997 Aug;22(3):190–9.
43. Corrigan P, Thompson V, Lambert D, et al. Perceptions of discrimination among persons with serious mental illness. Psychiatr Serv. 2003 Aug;54(8):1105–10.
44. Ren XS, Amick BC, Williams DR. Racial/ethnic disparities in health: the interplay between discrimination and socioeconomic status. Ethn Dis. 1999 Spring–Summer; 9(2):151–65.
45. Brondolo E, Kelly K, Coakley V, et al. The perceived ethnic discrimination questionnaire: development and preliminary validation of a community version. J App Soc Psychology. 2005;35(2):335–65.
46. Krieger N, Smith K, Naishadham D, et al. Experiences of discrimination: validity and reliability of a self-report measure for population health research on racism and health. Soc Sci Med. 2005 Oct;61(7):1576–96. Epub 2005 Apr 21.
47. Utsey SO. Development and validation of a short form of the Index of Race-Related Stress (IRRS)—brief version. Meas and Eval in Counsl and Dev. 1999;32:149–67.
48. Ryan AM, Gee GC, Laflamme DF. The association between self-reported discrimination, physical health and blood pressure: findings from African Americans, Black

immigrants, and Latino immigrants in New Hampshire. J Health Care Poor Underserved. 2006 May;17(2 Suppl):116–32.
49. Clark R, Gochett P. Interactive effects of perceived racism and coping responses predict a school-based assessment of blood pressure in Black youth. Ann Behav Med. 2006 Aug;32(1):1–9.
50. Clark R. Self-reported racism and social support predict blood pressure reactivity in Blacks. Ann Behav Med. 2003 Spring;25(2):127–36.
51. Eccles JS, Wong CA, Peck SC. Ethnicity as a social context for the development of African-American adolescents. J Sch Psychology. 2006;44:407–26.
52. Trivedi AN, Ayanian JZ. Perceived discrimination and use of preventive health services. J Gen Intern Med. 2006 Jun;21(6):553–8.
53. Williams DR. Racism and health: a research agenda. Ethn Dis. Ethn Dis. 1996 Winter–Spring;6(1–2):1–8.
54. Lazarus R, Folkman S. Stress, appraisal, and coping. New York: Springer, 1984.
55. Bishop GD. Health psychology, integrating mind and body. Needham Heights, MA: Allyn and Bacon, 1994.
56. Krieger N, Sydney S. Racial discrimination and blood pressure: the CARDIA Study of young black and white adults. Am J Public Health. 1996 Oct;86(10):1370–8.
57. Plous S. Understanding prejudice and discrimination. New York: McGraw Hill Companies, Inc., 2003.
58. Krieger N, Sydney S. Prevalence and health implications of anti-gay discrimination: a study of black and white women and men in the CARDIA cohort. Coronary Artery Risk Development in Young Adults. Int J Health Serv. 1997;27(1):157–76.
59. Meyer IH. Prejudice as stress: conceptual and measurement problems. Am J Public Health. 2003 Feb;93(2):262–5.
60. Ruggiero KM, Taylor DM. Why minority group members perceive or do not perceive the discrimination that confronts them: the role of self-esteem and perceived control. J Pers Soc Psychol. 1997;72(2):373–89.
61. Williams DR. Discrimination. San Francisco, CA: John D. and Catherine T. MacArthur Research Network on Socioeconomic Status and Health.
62. Baldwin JR, Day LE, Hecht ML. The structure(s) of racial attitudes among white college students. Intl J of Intercult Relat. 2000;24(5):553–77.
63. Williams DR. Race and health: basic questions, emerging directions. Ann Epidemiol. 1997;7:322–33.

Can Primary Care Visits Reduce Hospital Utilization Among Medicare Beneficiaries at the End of Life?

Andrea C. Kronman, MD, MSc[1], Arlene S. Ash, PhD[1], Karen M. Freund, MD, MPH[1], Amresh Hanchate, PhD[1], and Ezekiel J. Emanuel, MD, PhD[2]

[1]Women's Health and Health Care Research Units, Section of General Internal Medicine, Evans Department of Medicine, and the Women's Health Interdisciplinary Research Center, Boston University Medical Center, Boston, MA, USA; [2]Department of Clinical Bioethics, Warren Magnuson Clinical Center, National Institutes of Health, Bethesda, MD, USA.

BACKGROUND: Medical care at the end of life is often expensive and ineffective.

OBJECTIVE: To explore associations between primary care and hospital utilization at the end of life.

DESIGN: Retrospective analysis of Medicare data. We measured hospital utilization during the final 6 months of life and the number of primary care physician visits in the 12 preceding months. Multivariate cluster analysis adjusted for the effects of demographics, comorbidities, and geography in end-of-life healthcare utilization.

SUBJECTS: National random sample of 78,356 Medicare beneficiaries aged 66+ who died in 2001. Non-whites were over-sampled. All subjects with complete Medicare data for 18 months prior to death were retained, except for those in the End Stage Renal Disease program.

MEASUREMENTS: Hospital days, costs, in-hospital death, and presence of two types of preventable hospital admissions (Ambulatory Care Sensitive Conditions) during the final 6 months of life.

RESULTS: Sample characteristics: 38% had 0 primary care visits; 22%, 1–2; 19%, 3–5; 10%, 6–8; and 11%, 9+ visits. More primary care visits in the preceding year were associated with fewer hospital days at end of life (15.3 days for those with no primary care visits vs. 13.4 for those with ≥9 visits, P<0.001), lower costs ($24,400 vs. $23,400, P<0.05), less in-hospital death (44% vs. 40%, P<0.01), and fewer preventable hospitalizations for those with congestive heart failure (adjusted odds ratio, aOR=0.82, P<0.001) and chronic obstructive pulmonary disease (aOR=0.81, P=0.02).

CONCLUSIONS: Primary care visits in the preceding year are associated with less, and less costly, end-of-life hospital utilization. Increased primary care access for Medicare beneficiaries may decrease costs and improve quality at the end of life.

Received August 23, 2007
Revised January 30, 2008
Accepted April 8, 2008
Published online May 28, 2008

KEY WORDS: end-of-life care; health services research; primary care.
J Gen Intern Med 23(9):1330–5
DOI: 10.1007/s11606-008-0638-5

BACKGROUND

Medical treatments for the 6% of Medicare beneficiaries who die each year consume almost 30% of Medicare expenditures[1,2]. In addition, the quality of end-of-life health care is often poor[3]. Problems include late referrals to hospice, undertreatment of pain, over-treatment with unwanted or ineffective procedures, poor communications regarding prognosis and treatment preferences, and more in-hospital deaths that are inconsistent with stated preferences[1,3,4].

Several strategies have been proposed to reduce end-of-life health-care costs while improving quality, including increased use of advance directives and earlier referral to hospice. However, despite some evidence regarding improved quality, neither strategy clearly reduces costs for elderly Medicare beneficiaries[4–7].

Continuity of care has been associated with patients and their families experiencing a "higher quality death"[8], with fewer emergency department visits for cancer patients[9], increased patient satisfaction, increased adherence to recommended care, and less duplicate testing[10–13]. Although some studies have shown that continuity of care with a primary care physician has been associated with reduced healthcare costs and utilization in some patients[11,14], it remains unclear whether primary care leads to more appropriate care at the end of life. More care at the end of life by a primary care physician could enhance quality and reduce costs, since the provider may have more opportunities to prevent medical complications, discuss patient preferences, and coordinate home palliative care.

To assess the impact of primary care on end-of-life healthcare utilization, we explored whether more primary care visits were associated with key outcomes during the last 6 months of life: (1) fewer hospitalized days, (2) fewer in-hospital deaths, (3) fewer preventable hospital admissions, and (4) lower costs.

METHODS

Data Source

We examined primary care physician visits provided during the preceding 12 months ("pre-period") as a predictor of hospital

use and costs in the last 6 months of life. We used a randomly sampled population of 116,318 Medicare beneficiaries aged 66 or older who died in the last 6 months of 2001. Non-Whites were over-sampled, because the study population had been constructed to focus on end-of-life health-care disparities. To ensure completeness and comparability of healthcare utilization records prior to death, we examined only those in our study sample with complete data during their final 18 months. We excluded people not continuously enrolled in the Medicare parts A and B traditional fee-for-service program, who could not be matched to the National Death Index, and who were enrolled in the End Stage Renal Disease (ESRD) program. This left a final analytical sample of 78,353.

Age, sex, race, and zip code of residence were obtained from the Medicare denominator file, using the Medicare racial/ethnic categories of White, Black, Hispanic, and Other (for those of Asian, North American Native, and other or unknown races and ethnicities). A Medicaid indicator in the Medicare file was used as a proxy for low economic status. To adjust for nursing home status, and since we could not determine nursing home residence directly, we coded "any nursing home use" for people who used a Medicare-reimbursed skilled nursing facility (SNF) in the pre-period. A summary comorbidity measure was determined using DxCG's prospective relative risk score (DCG version 6.1 for Windows), derived from the presence of ICD-9-CM diagnosis codes from inpatient and outpatient encounters in Medicare's utilization files. These encounters include all physician visits, hospital care, and nursing home care, but not codes used for diagnostic tests. The score is calibrated to associate 1.0 with average expected expenditures in the following year among all Medicare beneficiaries observed during routine 12-month periods[15].

We used Berenson-Eggers-Type-of-Service (BETOS) codes in the Medicare Carrier file [http://www.cms.hhs.gov/HCPCSReleaseCodeSets/20_BETOS.asp] to identify outpatient visits in a nursing facility or office. We used the Medicare HCFA specialty codes to define a visit to an internist (11), geriatrician (38), or family practitioner (08), as "primary care"[16,17]. We used the number of such visits in the 12-month "pre-period" prior to the final 6 months of life to form five primary care groups: 0, 1–2, 3–5, 6–8, and ≥9.

Outcomes

We studied four outcomes during the last 6 months of life: (1) number of inpatient days [obtained from the Medicare Provider Analysis and Review (MedPAR) files], (2) in-hospital death (from the National Death Index), (3) total costs paid by Medicare (from the MedPAR, Carrier, Durable Medical Equipment, Hospice, and Outpatient files), and (4) any hospital admission for each of two chronic Ambulatory Care Sensitive Conditions (ACSC), congestive heart failure (CHF), and chronic obstructive pulmonary disease (COPD). Admissions for CHF and COPD, which are common chronic diseases in the elderly, can often be prevented by appropriate primary care[18,19].

Statistical Analysis

We used bivariate analyses (chi-square for categorical variables and analysis of variance for continuous variables) to identify differences in end-of-life utilization and costs across the primary-care groups. Due to large samples, almost all comparisons were highly statistically significant. STATA version 9.1 was used for all analyses.

To account for geographic differences, we used multivariable cluster analysis, specifically, fixed effects difference regression. This accounts for both measured and unmeasured healthcare supply and labor factors, which vary by geographic location, by only contrasting each outcome for beneficiaries residing in the same geographic area. We mapped each beneficiary's zip code of residence into the two Dartmouth Atlas-based geographic area units characterized by healthcare utilization patterns: the "hospital referral regions" (HRRs) and "hospital service areas" (HSAs). The United States is divided into approximately 300 HRRs and 3,000 HSAs. The five "primary-care visit groups" based on number of such visits during the pre-period were our key predictors, while adjusting for other factors known to affect health utilization and outcomes: age[20], sex[21], race [22], receipt of Medicaid, nursing home use, comorbidity[1,3,23,24], and place of residence[25–27] by HSA. We calculated each risk-adjusted, expected outcome for a primary-care visit group by using its coefficient and the mean values for each of the other covariates in the equation predicting the outcome.

Since sicker people visit doctors more often, the primary care visit group is highly confounded with pre-period morbidity. We additionally examined our data within comorbidity score quartiles to examine the possibly differential effect of primary-care visit frequency on sicker patients.

We examined the presence of end-of-life hospitalizations for CHF or COPD among patients with those conditions. We identified people with COPD (ICD-9-CM codes for chronic bronchitis 491.xx, emphysema 492.xx, asthma 493.xx, and COPD NOS 496.xx) or CHF (428.xx) if they had at least one clinical encounter with a diagnostic code for these conditions during the pre-period. We used logistic models to predict the likelihood of a hospitalization for each condition, adjusting for age, sex, race, Medicaid receipt, nursing home use, and total comorbidity burden (DCG score). However, with relatively rare dichotomous outcomes (only 8–17% of those with CHF or COPD were hospitalized for these conditions), it was not feasible to also account for geography using 3,000 regional fixed-effect clusters (HSAs). Recognizing the importance of geography, we verified that findings from the logistic models without geographic adjustment were consistent with a linear regression analysis of the same data, using hospital referral regions (HRRs) as the geographic cluster unit. Noting that pre-period primary visits were strongly associated with death in a nursing home (ranging from 18% for those with 0 visits to 43% for those with greater than 9 visits), we repeated the modeling including interactions between the pre-period nursing home use indicator and the primary care visit groups. We also performed several sensitivity analyses, examining separately those who did and did not use a skilled nursing facility (SNF) in the pre-period, and both controlling for hospice use, and removing hospice patients, from analyses.

RESULTS

Among 78,356 Medicare decedents in our sample, 56% were female; 40%, White; 36%, Black; and 11%, Hispanic. The mean age at death was 81 years (range 66–98). In the 12-month pre-period prior to the final 6 months of life, 32% received

Medicaid, 13% had Medicare-reimbursed nursing home care, and less than 2% were enrolled in hospice. Also, 38% had 0 primary care visits, 22% had 1–2, 19% had 3–5, 10% had 6–8, and 11% had 9 or more visits. The following were associated with less primary care utilization: younger age, Black race, male sex, no Medicaid, no nursing home use, fewer hospital admissions, and less comorbidity (Table 1).

In the final 6 months of life, 24% of the population used hospice. Death occurred in a hospital for 43% and a nursing home for 25%. The average number of hospital days was 15.1, average costs were $24,800, and 17% of those with CHF and 8% of those with COPD had at least ine admission for the respective ACSC (Table 1).

More primary care visits in the pre-period were associated with reduced hospital days, in-hospital deaths, cost, and preventable hospital admissions. After adjusting for age, sex, race, Medicaid, nursing home use, comorbidities, and geographic location, expected total hospital days in the last 6 months of life varied with the number of primary-care visits in the pre-period as follows (Table 2): 15.3 days for decedents with 0 primary-care visits; 15.9 days for 1–2 visits; 14.2 days for 6–8 visits; and 13.4 days for 9 or more visits (P<0.001).The association was even greater among those who used SNF care in the pre-period (Table 3).

After adjusting for age, sex, race, Medicaid, nursing home use, comorbidities, and place of residence, in-hospital deaths occurred as follows: 43.9% of decedents with 0 primary-care visits, 43.8% of those with 1–2 visits, 43.1% with 3–5 visits, 39.5% with 6–8 visits, and 39.2% of those with ≥9 visits (Table 2). While hospice use was associated with site of death, sensitivity analyses that either controlled for hospice use or removed hospice patients did not notably alter the association between primary care visits and in-hospital death (data not shown).

Adjusting for the same factors, more primary-care visits were associated with reduced total Medicare expenditures at the end of life (Table 2). Among decedents with 0 primary-care visits in the pre-period, total costs in the last 6 months were $24,449, compared to $26,026 for decedents with 1–2 primary care visits, $25,572 for decedents with 3–5, $24,005 for decedents with 6–8 primary care visits, and $23,345 for decedents with ≥9 primary care visits.

Among Medicare beneficiaries with diagnoses of congestive heart failure (CHF) or chronic obstructive pulmonary disease (COPD), those who had more primary care visits in the pre-period were less likely to be hospitalized for these conditions during the last 6 months of life (Table 3). Those with ≥9 primary care visits in the 12 months preceding the end-of-life period were significantly less likely to be hospitalized for CHF (odds ratio=0.82, 95% CI 0.74—0.92), and COPD (odds ratio= 0.81, 95% CI 0.68–0.97) compared to those with fewer visits. These significant associations between more primary care visits and the main outcome of hospital days were magnified for the sickest 25% of patients. Repeating the multivariable

Table 1. Decedent Characteristics by Numbers of Prior Primary Care Visits* †

Pre-period Characteristics†	Primary Care Visit Groups					
	Total	0	1–2	3–5	6–8	≥9
N	78,356	29,557	17,181	15,112	7,952	8,554
%	100	38	22	19	10	11
Mean age (SD)	80.9 (8.1)	80.1 (8.2)	80.8 (8.1)	81.2 (8.0)	81.6 (8.0)	82.5 (8.1)
Women, %	56	51	57	59	61	63
White, %	40	36	40	42	44	41
Black, %	36	41	36	32	30	31
Hispanic, %	11	11	11	11	10	10
Other, %	14	13	13	14	16	17
Medicaid, %	32	29	31	32	36	42
Nursing home use,%	13	6	12	16	20	27
Mean comorbidity risk score (SD)‖	2.2 (1.7)	1.7 (1.5)	2.1 (1.6)	2.4 (1.6)	2.8 (1.7)	3.3 (1.9)
Mean number of hospital admissions (SD)		0.7 (1.3)	0.9 (1.4)	1.0 (1.5)	1.1 (1.6)	1.3 (1.7)
Diagnoses:						
CHF, % §	32	23	29	36	43	48
COPD, % §	26	19	25	30	33	35
End-of-life utilization‡						
Mean total hospital days (SD)	15.1 (20.2)	14.5 (20.2)	15.5 (19.8)	15.6 (20.2)	15.0 (20.1)	15.8 (21.1)
Mean total costs in $1,000 (SD)	24.8 (30.9)	23.4 (29.8)	25.1 (30.2)	25.7 (30.3)	25.2 (31.2)	27.4 (36.3)
Hospice, %	24	22	25	25	24	22
Place of death						
In-hospital, %	43	45	44	42	39	39
Nursing home, %	25	18	22	28	36	43
Residence, %	21	24	22	19	16	10
ACSC hospitalization among those with:						
CHF, % §	17	18	18	17	16	15
COPD, %§	8	10	9	8	7	7

**All P<0.001*

†Pre-period, months 18–7 before death

‡End-of-life utilization, months 6–date of death, unadjusted

§CHF, congestive heart failure, COPD, chronic obstructive pulmonary disease

‖Comorbidity risk score = relative risk from DxCG's prospective risk adjustment software, which organizes ICD-9-CM diagnosis codes from the Medicare utilization files, assigns weights to them, and summarizes their expected impact on future expenditures via a relative risk score.

Table 2. Healthcare Utilization[§] and Percentage of Deaths[§] in Hospital by Level of Prior Primary Care Use [||] (N=78,356)

Number of Primary Care Visits	Total Hospital Days (95% CI)	Total Costs $1,000 (95% CI)	In-hospital Death % Population (95% CI)
0	15.3 (15.0, 15.5)	24.5 (24.1, 24.8)	43.9 (43.3, 44.5)
1–2	15.9 (15.6, 16.2) [‡]	26.0 (25.6, 26.5) *	43.8 (43.1, 44.5)
3–5	15.4 (15.1, 15.8)	25.7 (25.1, 26.0) *	43.1 (42.3, 43.9) [†]
6–8	14.2 (13.7, 14.6) [‡]	24.0 (23.3, 24.7)	39.5 (38.4, 40.1) [†]
≥9	13.4 (12.9, 13.8) [‡]	23.4 (22.7, 24.0) *	39.2 (38.1, 40.4) [†]

*$P<0.05$, reference=0 primary-care visits
[†]$P<0.01$, reference=0 primary-care visits
[‡]$P<0.001$, reference=0 primary-care visits
[§]*Outcomes measured during final 6 months of life and adjusted for age, sex, race, Medicaid, nursing home use, comorbidity, and geographic variation (hospital service area)*
[||]*Primary-care visits measured during pre-period, months 18-7 before death*

analyses using comorbidity score quartiles, we found the sickest decedents with ≥9 primary care visits had an average of two fewer hospital days compared to those with no primary care visits (Table 3).

DISCUSSION

Frequent primary care visits were associated with four key end-of-life care outcomes: fewer days hospitalized, fewer preventable hospital admissions, fewer in-hospital deaths, and lower total costs.

Few interventions have been shown to influence end-of-life care either by improving quality or reducing costs. Although hospice and advance directives can improve patient self-efficacy at the end of life, neither clearly reduces costs[4,5,7,28,29]. Thus, the association of visits to primary care physicians with substantial reductions in costs and utilization at the end of life is especially notable. Although we cannot conclude from our cross-sectional analysis that more primary care visits *cause* lower utilization, primary care may substitute outpatient visits for more costly hospitalizations of patients with complex medical conditions. Although previously shown for specific chronic diseases[30], our study is the first to find this association at the end of life. A recent systematic review found that hospitalization of nursing home residents is determined by many factors, including sociodemographics, individual preferences, provider preferences, and economics of the particular healthcare system[31]. Our study suggests that fewer prior primary care visits are yet another determinant of hospitalization for Medicare beneficiaries.

There may be a threshold effect, because only at six or more visits was end-of-life utilization lower. Higher utilization and costs incurred by those with 1–5 visits compared to those with 0 visits could be due to patients with relatively high morbidity receiving too few primary care visits. In addition, patients with 0 visits had healthier beneficiaries who likely required few healthcare services; the healthiest patients probably do not benefit from frequent primary care visits. However, among the sickest Medicare beneficiaries, frequent primary care visits were associated with a 9% reduction in hospital days.

Our findings differ from those of Weinberger et al.[32], who found that increased primary care was associated with a higher hospital re-admission rate. It is possible that for the select group of severely ill veterans in this study, more primary care led to more hospital re-admissions because the patients were prematurely discharged from the hospital, and their clinical decompensation was appropriately evaluated and triaged by the primary care team. In contrast, our study population is 50-fold larger and a more representative sample of the national Medicare population. In addition, rather than focusing on re-admissions, we measured total hospital admissions and days.

By using fixed effect regression analysis with geographic clustering, we adjusted for both measurable and unmeasurable geographic factors. Previous studies have shown the importance of local characteristics of the health-care system in rates of preventable hospitalizations for Ambulatory Care

Table 3. Hospital Utilization[§] and ACSC[||] Admissions within Selected Patient Cohorts by Level of Prior Primary Care Use[¶]

Number of Primary Care Visits	Non-SNF Users Hospital Days (95% CI) N=68,170	SNF Users Hospital Days (95% CI) N=10,186	Sickest Quartile Hospital Days (95% CI) N=19,589	CHF Admission Odds Ratio (95% CI) N=24,856	COPD Admission Odds Ratio (95% CI) N=20,161
0	15.4 (15.1,15.6)	14.5 (13.5,15.4)	22.7 (21.6, 23.9)[‡]	Reference	Reference
1–2	16.2 (15.9,16.5)[†]	13.9 (13.0,14.8)	21.9 (20.8, 23.0)[‡]	1.00 (0.93,1.12)	0.96 (0.84,1.10)
3–5	15.7 (15.4,16.1)	13.8 (12.9,14.6)[†]	21.1 (19.8, 22.4)[‡]	0.98 (0.89,1.08)	0.85 (0.74,0.98)*
6–8	14.4 (13.9,14.9)[†]	12.9 (11.9,13.9)[†]	20.5 (19.3, 21.7)[‡]	0.88 (0.79,0.99)*	0.75 (0.63,0.90)*
≥9	13.8 (13.3,14.3)[†]	11.6 (10.8,12.5)[†]	19.5 (18.8, 20.3)[‡]	0.82 (0.74,0.92)*	0.81 (0.68,0.97)*

*$P<0.05$, reference=0 primary-care visits
[†]$P<0.01$, reference=0 primary-care visits, no Skilled Nursing Facility (SNF) services
[‡]$P<0.01$, reference=0 primary-care visits, lowest quartile comorbidity (least sick), after population was first stratified into four quartiles by comorbidity risk score
[§] *Utilization measured during final 6 months of life, and adjusted for age, sex, race, Medicaid, nursing home use, comorbidity, geographic variation (hospital service area)*
[||]*ACSC, ambulatory care sensitive conditions: CHF, congestive heart failure; COPD, chronic obstructive lung disease. Admissions for ACSC measured during final 6 months of life and adjusted for age, sex, race, Medicaid, nursing home use, and comorbidity*
[¶]*Primary-care visits measured during pre-period, months 18-7 before death*

Sensitive Conditions[30,33] and hospital utilization at the end of life[26,27,34]. These local characteristics are more important than patient preferences in determining whether someone dies in a hospital at the end of life[21]. After controlling for geographic variations, we found that fewer previous primary care visits are also a determinant of in-hospital death. Our findings are robust to geographic area variations in healthcare use. With more visits, primary care physicians may be better able to elicit patients' preferences, resulting in fewer hospitalizations and unwanted in-hospital deaths.

The study has several limitations. Medicare claims data contain no direct information regarding beneficiary preferences, appropriateness of clinical treatment, or quality of care. These findings may not generalize to Medicare beneficiaries in the End Stage Renal Disease program, in managed care plans, or those without the optional Medicare part B coverage. Although we did not have clinical data on disease severity, the DCG comorbidity score, constructed from detailed data encoded in diagnoses recorded during all medical encounters, has been shown to accurately predict future utilization of a population with very different levels of future mortality and costs[15,24,35,36]. Primary care visits do not account for all forms of primary care. For example, nurse visits, telephone consultations, and primary care provided by specialists were not counted. The economic status of beneficiaries was only partially captured by Medicaid receipt as noted in Medicare's records. Also, since Medicare only covers the first 3 months of a nursing home stay after hospitalization, we could not distinguish long-term care residents from short-term skilled nursing facility (SNF) users in our dataset. However, our findings suggest that primary care visits reduce hospital utilization most profoundly within nursing homes. Almost half of the beneficiaries with >9 primary-care visits died in a nursing home, and the association of more primary-care visits with reduced end-of-life utilization was most striking among previous SNF users. Given these limitations, future studies could use the additional demographic information in the Medicare Minimum Dataset (MDS) to clarify the relationship between primary care visits and hospital utilization among long-term nursing home residents.

As concerns about the quality and costs of end-of-life care increase, our study suggests that providing more primary care to Medicare beneficiaries may improve the quality of end-of-life care while reducing time spent in the hospital and overall costs. In 2001, nine primary-care visits cost Medicare approximately $3,000, 9 days in the hospital cost Medicare approximately $11,000, and 533,000 fee-for-service Medicare beneficiaries died in the hospital[37]. Decreasing just 1 hospital day for each of these beneficiaries at the end of life could have saved millions of dollars. Future studies that incorporate Medicare's DRG reimbursement system, hospice, home services, and pharmaceutical costs are needed to validate the cost-effectiveness of enhanced primary care at the end of life.

To achieve greater primary care utilization by a growing population of elderly, the workforce of primary care providers must grow. However, the primary care workforce is diminishing due to many primary care physicians leaving practice and few young physicians entering primary care[38]. Thus, providing more primary care may require increased training opportunities for nurses and physicians, or altered incentives that make primary care provision a more attractive enterprise[33,38].

Acknowledgements: *This work was presented at the Academy-Health Annual Research Meeting in Orlando, FL, on 4 June 2006 and received the Mack Lipkin Sr. Associate Member Award for Outstanding Scientific Presentation at the National Meeting for the Society of General Internal Medicine in Los Angeles, CA, on 26 April 2006. The authors thank Drs. Jim Burgess and John Pagliaro for their support with manuscript preparation, and Jenn Fonda for support with computer programming. Dr. Kronman is supported by a Mentored Clinical Scientist Development Program Award (Building Interdisciplinary Careers in Women's Health Research) from the National Institutes of Health (K12 HD04344, Karen Freund, Principal Investigator). Dr. Kronman was supported by an Institutional National Research Service Award from the National Institutes of Health (T32 HP 10028, Rob Friedman, Principal Investigator) when this research was performed.*

Conflict of Interest: *None disclosed.*

Corresponding Author: *Andrea C. Kronman, MD, MSc; Women's Health and Health Care Research Units, Section of General Internal Medicine, Evans Department of Medicine, and the Women's Health Interdisciplinary Research Center, Boston University Medical Center, 801 Massachusetts Ave, Suite 470, Boston, MA 02118, USA (e-mail: Andrea.Kronman@bmc.org).*

REFERENCES

1. **Hogan C, Lunney J, Gabel J, Lynn J.** Medicare beneficiaries' costs of care in the last year of life. Health Aff. 2001;204188–95.
2. **Edwards C, DeHaven T.** War betweeen the generations: federal spending on the elderly set to explode. In: C. Institute, eds. Policy Analysis. Washington D.C.; 2003: 1–22.
3. Approaching Death: Improving Care at the End of Life. Washington D.C.: National Academy Press; 1997.
4. **Ezekiel EJ.** Cost savings at the end of life: What do the data show? JAMA. 1996;275241907–14.
5. **Campbell D, Lynn J, Louis T, Shugarman L.** Medicare program expenditures associated with hospice use. Ann Intern Med. 2004; 140:269–77.
6. **Degenholtz HB, Rhee Y, Arnold RM.** Brief communication: the relationship between having a living will and dying in place. Ann Intern Med. 2004;1412113–7.
7. SUPPORT Principal Investigators. A controlled trial to improve care for seriously ill hospitalized patients: the study to understand prognoses and preferences for outcomes and risks of treatments. JAMA. 1995;274:1591–8.
8. **Patrick D, Curtis R, Engelberg R, Nielsen E, McCown E.** Measuring and improving the quality of dying and death. Ann Intern Med. 2003;1392410–5.
9. **Burge F, Lawson B, Johnston G.** Family physician continuity of care and emergency department use in end-of-life cancer care. Med Care. 2003;418992–1001.
10. **Burge F, Lawson B, Johnston G, Cummings I.** Primary care continuity and location of death for those with cancer. J Palliat Med. 2003;66911–8.
11. **De Maeseneer JM, De Prins L, Gosset C, Heyerick J.** Provider continuity in family medicine: does it make a difference for total health care costs? Ann Fam Med. 2003;13144–8.
12. **Gill JM, Mainous AG 3rd.** The role of provider continuity in preventing hospitalizations. Arch Fam Med. 1998;74352–7.
13. **Saultz JW, Albedaiwi W.** Interpersonal continuity of care and patient satisfaction: a critical review. Ann Fam Med. 2004;25445–51.
14. **Cabana MD, Jee SH.** Does continuity of care improve patient outcomes? J Fam Pract. 2004;5312974–80.
15. **Ash AS, Ellis RP, Pope GC, et al.** Using diagnoses to describe populations and predict costs. Health Care Financ Rev. 2000;2137–28.
16. **Burns RB, McCarthy EP, Freund KM, et al.** Black women receive less mammography even with similar use of primary care. Ann Intern Med. 1996;125:173–82.
17. **Keating NL, Landrum MB, Ayanian JZ, et al.** The association of ambulatory care with breast cancer stage at diagnosis among Medicare beneficiaries. J Gen Intern Med. 2005;20:38–44.

18. **Ricketts TC, Randolph R, Howard HA, Pathman D, Carey T.** Hospitalization rates as indicators of access to primary care. Health & Place. 2001;7127–38.
19. Agency for Healthcare Research and Quality. Refinement of the HCUP Quality Indicators. 2001 Summary, Technical Review Number 4. Available at: http://www.qualityindicators.ahrq.gov/downloads/technical/qi_technical_summary.pdf. Accessed March 12, 2008.
20. **Norman Levinsky M, Wei Yu P, Arlene Ash P, et al.** Influence of age on medicare expenditures and medical care in the last year of life. JAMA. 2001;286111349–55.
21. **Valentin A, Jordan B, Lang T, Hiesmayr M, Metnitz PGH.** Gender-related differences in intensive care: A multiple-center cohort study of therapeutic interventions and outcome in critically ill patients. Crit Care Med. 2003;3171901–7.
22. **Degenholtz HB, Thomas SB, Miller MJ.** Race and the intensive care unit: disparities and preferences for end-of-life care. Crit Care Med. 2003;315 SupplS373–8.
23. **Wolff JL, Starfield B, Anderson G.** Prevalence, expenditures, and complications of multiple chronic conditions in the elderly. Arch Intern Med. 2002;162202269–76.
24. **Ash AS, Posner MA, Speckman J, Franco S, Yacht AC, Bramwell L.** Using claims data to examine mortality trends following hospitalization for heart attack in Medicare. Health Serv Res. 2003;3851253–62.
25. **Pritchard RS, Fisher ES, Teno JM, et al.** Influence of patient preferences and local health system characteristics on the place of death. SUPPORT investigators. Study to understand prognoses and preferences for risks and outcomes of treatment. J Am Geriatr Soc. 1998;46101242–50.
26. **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. 2004;(Suppl Web Exclusive):VAR33–44.
27. **Wennberg JE, Fisher ES, Stukel TA, Skinner JS, Sharp SM, Bronner KK.** Use of hospitals, physician visits, and hospice care during last six months of life among cohorts loyal to highly respected hospitals in the United States. BMJ. 2004;3287440607.
28. **Emanuel E, Emanuel L.** The economics of dying–the illusion of cost savings at the end of life. N Engl J Med. 1994;3308540–4.
29. **Shugarman LR, Campbell DE, Bird CE, Gabel J, Louis TA, Lynn J.** Differences in medicare expenditures during the last 3 years of life. J Gen Intern Med. 2004;192127–35.
30. **Bodenheimer T, Wagner EH, Grumbach K.** Improving primary care for patients with chronic illness: the chronic care model, Part 2. JAMA. 2002;288151909–14.
31. **Grabowski DC, Stewart KA, Broderick SM, et al.** Predictors of nursing home hospitalization: a review of the literature. Med Care Res Rev. 2008;6513–39.
32. **Weinberger M, Oddone EZ, Henderson WG.** Does increased access to primary care reduce hospital readmissions? Veterans affairs cooperative study group on primary care and hospital readmission. N Engl J Med. 1996;334221441–7.
33. **Bodenheimer T.** Primary care–will it survive? N Engl J Med. 2006;3559861–4.
34. **Fisher ES, Wennberg JE, Stukel TA, et al.** Associations among hospital capacity, utilization, and mortality of US Medicare beneficiaries, controlling for sociodemographic factors. Health Serv Res. 34(6):1351–62.
35. **Zhao Y, Ash AS, Ellis RP, Slaughter JP.** Disease burden profiles: an emerging tool for managing managed care. Health Care Manage Sci. 2002;53211–9.
36. **Kronman A, Hanchate A, Ash AS.** Improving risk adjustment for illness burden. J Gen Intern Med. 2007;22(s1). Abstract.
37. Health Care Financing Review: Medicare and Medicaid Statistical Supplement. US Dept. of Health and Human Services, Centers for Medicare and Medicaid Services, Office of Research, Development and Information, 2003.
38. Statement for the Record for Senate HELP Committee Hearing on Addressing Healthcare Workforce Issues for the Future (12-Feb-08): A Report from the American College of Physicians; 2008. Available at: http://www.acponline.org/advocacy/where_we_stand/workforce/pc_workforce.pdf Accessed March 12, 2008.

Implications of Comorbidity on Costs for Patients With Alzheimer Disease

Tzu-Chun Kuo, PhD, Yang Zhao, PhD,† Sharada Weir, DPhil,‡§ Marilyn Schlein Kramer, MBA,* and Arlene S. Ash, PhD*‡§*

Background: No prior studies have used a comprehensive clinical classification system to examine the effect of differences in overall illness burden and the presence of other diseases on costs for patients with Alzheimer disease (AD) when compared with demographically matched nondemented controls.

Study Design: Of a total of 627,775 enrollees who were eligible for medical and pharmacy benefits for 2003 and 2004 in the MarketScan Medicare Supplemental and Coordination of Benefits Database, we found 25,109 AD patients. For each case, 3 demographically matched nondemented controls were selected using propensity scores. Applying the diagnostic cost groups (DCGs) model to all enrollees, 2003 diagnoses were used to estimate prospective relative risk scores (RRSs) that predict 2004 costs from all illness other than AD. RRSs were then used to control for illness burden to estimate AD's independent effect on costs.

Results: Compared with the control group, the AD cohort has more comorbid conditions (8.1 vs. 6.5) and higher illness burden (1.23 vs. 1.04). Individuals with AD are more likely to have mental health conditions, neurologic conditions, cognitive disorders, cerebrovascular disease, diabetes with acute complications, and injuries. Annual costs for AD patients are $3567 (34%) higher than for controls. Excess costs attributable to AD, after controlling for non-AD illness burden, are estimated at $2307 per year with outpatient pharmacy being the key driver ($1711 in excess costs).

Conclusions: AD patients are sicker and more expensive than demographically matched controls. Even after adjusting for differences in illness burden, costs remain higher for AD patients.

Key Words: Alzheimer disease, comorbidity, illness burden, diagnostic cost group

(*Med Care* 2008;46: 839–846)

From the *Research Division, DxCG, Inc., Boston, Massachusetts; ‡Center for Health Policy and Research and Department of Family Medicine and Community Health, University of Massachusetts Medical School, Boston, Massachusetts; §School of Medicine, Boston University, Boston, Massachusetts; and †Eli Lilly and Company, Indianapolis, Indiana.

Supported by Eli Lilly and Company, Indianapolis, IN.

The statements contained in this paper are solely those of the authors and no endorsement by Eli Lilly and Company and DxCG, Inc. should be inferred or implied.

The study protocol was approved by Eli Lilly and Company, and DxCG, Inc.

Reprints: Tzu-Chun Kuo, PhD, Research Division, DxCG, Inc., 99 Summer Street, Suite 520, Boston, MA 02110. E-mail: tkuo@dxcg.com.

ISSN: 0025-7079/08/4608-0839

Alzheimer disease (AD) is a devastating neurodegenerative disease affecting 1 in 10 individuals over 65 and nearly half of those over 85.[1,2] As the population ages, the prevalence of AD will increase. By 2050, the worldwide number with AD is expected to grow from 26 million in 2006 to nearly quadruple by 2050,[3] with daunting implications for costs.[4,5] Examining the effect of comorbid conditions on costs should inform better management of AD patients.

Prior research has found that the medical care costs for AD patients are not only large and increasing but that their costs are substantially higher than for average Medicare beneficiaries in the United States.[6–19] Failure to understand and recognize these excess costs can result in inappropriate resource allocations and unfair efficiency assessments for providers of patients with AD. Prior studies have typically used the Charlson Comorbidity Index,[20] which summarizes the combined impact of 19 medical conditions, to examine the effect of comorbidities on costs in AD.[8–14] The Charlson Index includes no psychiatric comorbidities, urinary tract infection, pneumonia, or hip fracture; all of which are common among individuals with AD. No previous study has explicitly examined a comprehensive disease profile of AD patients and its impact on the total medical cost of AD patients.

Many AD cost studies have used Medicare administrative records with no information on outpatient prescription drugs, an important category of AD spending.[6–9,11,16–19] Also, most previous studies have examined data collected before 2000. In contrast we use a comprehensive claims database from 2003 to 2004 with complete medical and pharmacy claims. We also use a comprehensive clinical classification system to identify all comorbidities of AD, to summarize their collective impact in a prospective relative risk score (RRS), and to compare total medical costs for patients with AD with a demographically matched, nondemented control group. We use the RRS to measure "excess costs" in the AD patients compared with controls, to examine characteristics of the most and least expensive AD patients, and to test the hypotheses that AD patients incur higher and more predictable medical costs, and that they have more and different comorbidities, than demographically similar controls. Finally, we investigate the extent to which a greater burden of other illnesses explains the elevated costs of patients with AD. This information will help health plans develop and evaluate effective management protocols for AD.

METHODS

Data Source

Data were drawn from the MarketScan Medicare Supplemental and Coordination of Benefits (COB) Database for 2003 and 2004. The data captured person-specific clinical utilization, expenditures, and enrollment across inpatient, outpatient, and prescription drugs. These data contain claims of Medicare beneficiaries with employer-sponsored supplemental coverage, with both Medicare- and supplemental-insurance paid costs, including coinsurance and deductibles.

Sample Selection

The eligible population included those aged 65 or older as of January 1, 2003 who were eligible for medical and pharmacy benefits for all 12 months of 2003 and at least 1 month in 2004 (n = 627,775).

There are many types of non-Alzheimer's dementia, such as vascular dementia, Parkinson disease, Lewy body dementia, and frontotemporal dementia. The AD patient population accounts for 50–70% of all dementias and is the most homogeneous dementia.[21] In addition, AD is the only type of dementia with approved pharmacologic treatment and its pathology is the best understood. Because AD patients may be misdiagnosed with other dementias, we chose to make a clean comparison by comparing those diagnosed with AD with demographically similar controls with no evidence of dementia.

The AD cohort was selected according to clinical profiles and drug information, which included those who had at least 1 claim with ICD-9-CM code 331.0 (Alzheimer disease)[22] in any position on a physician/clinically trained professional or facility claim with a "from date of service" in 2003 and/or at least one pharmacy claim for Tacrine, Donepezil, Galantamine, Rivastigmine, or Memintine with a "filled date" in 2003. Using these criteria, we identified 25,109 individuals with AD in our data. In the nondemented control population, we excluded not only individuals with AD (as defined above), but also individuals with claims with any diagnosis of dementia (290.xx, 294.1x, 294.8x, 331.1x–331.9x, 797.xx) in any position on a physician/clinically trained professional or facility claim with a "from date of service" or a pharmacy claim for any of the drugs listed above with a "fill date" in either 2003 or 2004.

To select a nondemented control cohort that was demographically comparable with the AD cohort, we first used logistic regression to estimate the propensity for (that is, the probability of) being in the AD cohort as a function of 4 factors: age on January 1, 2003, sex, US geographic region (North East, North Central, West, South, and Unknown), and eligibility in 2004 (1–5, 6–11, and 12 month categories). We then partitioned the combined AD and nondemented population into 20.5%-quantiles of increasing propensity score. Within each quantile, we took all of the AD cases and a random sample from the nondemented control group. The size of MarketScan data set allowed us to select 3 people in the nondemented control group for every person in the AD cohort. Thus, the control cohort contained 75,327 people.

Illness Burden Measure

We implemented diagnostic cost groups (DCGs) models using DxCGs RiskSmart software to measure illness burden and account for the effect of illness burden on cost.[23] DCG models create comprehensive clinical profiles of each individual in the sample based on all diagnoses recorded on claims (excluding laboratory and radiology claims and other services without a face-to-face clinician encounter). ICD-9-CM diagnosis codes are grouped into 781 homogeneous clinical groups, called DxGroups. The DxGroups are mapped into 184 condition categories (CCs) based on the clinical body system and relative resource use. The CCs in turn are organized into 30 aggregated condition categories (ACCs) based on body system and condition, with CCs within a given ACC (eg, diabetes, heart disease, or cancer) being arranged in hierarchies so that only the most severe manifestation of each distinct type of condition is credited in a "hierarchicalized" vector of information that summarizes the conditions present (HCCs). Although the DxGroups are useful for "drilling down" to examine highly specific clinical entities, and the ACCs are mainly useful for high-level reporting, it is HCCs, age and sex that are used to predict outcomes, such as next year's costs, making DCG models robust to many variations in practices for coding diagnoses in administrative claims data.[24–31]

First, we used 2003 diagnoses (as grouped into the DCG clinical groupings) and demographics (age/sex) to predict 2004 costs for the entire population of 627,775 eligible individuals. This prospective framework is useful for segmenting populations based on future costs or utilization. We quantified "illness burden" for each individual as an RRS, which is his or her predicted cost divided by the mean cost of the full group of 627,775 people. Thus, the mean RRS for the full population is 1.00, and an individual with a RRS of, say, 1.20 is expected—based on illness burden—to cost 20% more than the population average. For this study, we "zeroed out"—that is, did not use—the CC for AD in calculating relative risk scores, so that 2 people of the same age and sex and identical disease profiles with the exception of AD are predicted to have the same cost. Thus, risk-adjusted differences in costs between the AD cohort and the control group can be exclusively attributed to the presence of AD.

Analysis

We used standard descriptive statistics to compare the costs of the AD and control cohorts, and the coefficient of variation (CV), defined as a population mean divided by its standard deviation and expressed as a percent, to compare cost variability. We calculated odds ratios to examine differences between the cohorts in the prevalence of other diseases. Weighted least squares (WLS) regressions were used to assess the impact of AD on total cost and cost by claim type (ie, inpatient, outpatient, and pharmacy), with the weights being fraction of the year eligible in 2004. We estimated WLS models of the following structure on the combined AD plus control cohorts:

$$Y = \alpha + \beta * (\text{AD Cohort}) + \gamma_1 * (\text{RRS}) + \gamma_2 * (\text{RRS}^2),$$

where Y is an annualized cost (for a type of service, or total), AD cohort is an indicator variable for being in the AD cohort and RRS is the prospective relative risk score as described above. β is the incremental cost associated with having AD.

RESULTS

Descriptive Statistics

As shown in Table 1, the AD cohort (n = 25,109) is on average older than the total nondemented population from which the control cohort was drawn (mean age 80.1 year vs. 74.2; 6.7% aged 90 or older vs. 2.3%); it has more females (61.6% vs. 56.2%), fewer people with a full year (12 months) of eligibility in 2004 (85.2% vs. 94.6%), and a higher burden of non-AD illness (mean RRS of 1.23 vs. 0.98). Thus, based on differences in age and other comorbidities alone, and ignoring the diagnosis of AD, we expect people with AD to cost quite a bit more than the nondemented population.

Propensity score matching yielded 75,327 controls that are similar to the AD cohort on age, sex, region, and eligibility. For example, the controls average age is 80.1 year, they are 62.5% female, and 87.6% are eligible for all 12 months in 2004. However, the matching purposefully allowed illness burden to reflect its natural distribution in the nondemented controls, who are older than the nondemented population as a whole, so we expect a higher risk score than 0.98. Because the matching does not enforce any particular relationship between the risk scores of the AD and control cohorts, it is meaningful that the control cohorts mean RRS is only 1.04, far less than the AD cohorts mean RRS of 1.23. Thus, based on a higher burden of non-AD illness burden alone, we expect the AD cohort to cost about 18% more than the controls.

Cost Variation

As seen in Table 2, AD patients are 34% more expensive than controls ($13,936 compared with $10,369 in 2004). They also cost more within each spending category (inpatient care cost, outpatient care cost, and outpatient pharmacy cost). As we will quantify later, part of the higher cost of AD is attributable to higher non-AD illness burden. Total cost and cost by spending category for the AD patients exhibit less variation (that is, smaller standard deviations) than costs for the control cohort. Together, higher costs and lower standard deviations lead to markedly lower CVs in the AD cohort. For example, the CV for total cost is 181 for the AD cohort versus 324 for controls.

Within each cohort, we examined costs by category and in total for the sickest 5% and 20% of the individuals—that is, those with the highest prospective RRSs—and compared

TABLE 1. Population Characteristics by Cohort

		Full Population			
	Full Population	AD Cohort	Non-AD Dementia	Non-Demented Control Population	Control Cohort*
Observations	627,775	25,109	33,787	568,879	75,327
Age as of January 1, 2003 (yrs)					
Mean	74.7	80.1	79.2	74.2	80.1
65–69 (%)	27.3	6.1	9.7	29.3	6.1
70–74 (%)	26.3	13.9	17.2	27.3	13.7
75–79 (%)	22.0	25.6	24.6	21.7	26.1
80–84 (%)	14.9	29.2	24.7	13.6	28.5
85–89 (%)	6.9	18.5	16.1	5.8	17.8
90+ (%)	2.7	6.7	7.7	2.3	7.8
Sex					
Male (%)	43.5	38.4	40.5	43.9	37.5
Female (%)	56.6	61.6	59.5	56.2	62.5
Region					
Northeast (%)	12.9	11.9	11.7	13.0	14.3
North central (%)	34.0	34.7	36.6	33.8	34.2
South (%)	34.9	36.6	34.3	34.8	31.9
West (%)	18.1	16.7	17.2	18.2	19.3
Unknown (%)	0.3	0.2	0.2	0.3	0.3
Eligibility months in 2004					
Mean months	11.6	11.1	11.3	11.7	11.2
1–5 months (%)	2.6	7.0	5.0	2.2	5.9
6–11 months (%)	3.6	7.8	7.5	3.2	6.5
12 months (%)	93.9	85.2	87.5	94.6	87.6
2003 morbidity indicators					
Mean illness burden score†	1.00	1.23	1.41	0.98	1.04
Mean number of condition categories	6.41	8.12	9.37	6.16	6.46

*Subset of the nondemented control population using propensity score based on age, sex, region, and 2004 eligibility months.
†Prospective and based on DCG Model with AD indicators zeroed out and normalized to the full population.

TABLE 2. 2004 Cost for AD and Control Cohorts, by Type of Service

	AD Cohort			Control Cohort			Risk-Adjusted Difference†	P‡
	Mean	SD	CV*	Mean	SD	CV*		
Inpatient care	$5111	21,396	419	$4024	30,243	751	$719	<0.0001
Outpatient care	$4769	9373	197	$4176	10,170	244	−$123	0.0489
Outpatient pharmacy	$4056	3109	77	$2169	3782	174	$1711	<0.001
Total	$13,936	25,214	181	$10,369	33,561	324	$2307	<0.001

*CV = 100 × SD/mean.

†The risk-adjusted difference is calculated as $\beta 1$ from a weighted least square regression of the form $Y = \alpha + \beta * (\text{AD Cohort}) + \gamma_1 * (\text{RRS}) + \gamma_2 * (\text{RRS}^2)$, where Y is total cost, AD indicates being in the AD Cohort and RRS is the DCG-based prospective relative risk score with AD indicators zeroed out.

‡The t test on the estimated $\beta 1$ coefficient, in the regression described below, for risk-adjusted cost.

TABLE 3. 2004 Cost by Quantiles of Predicted Cost, Within AD and Control Cohorts

	AD Cohort	Control Cohort	Risk-Adjusted Difference†	P‡
50% of entire cohort with lowest predicted cost				
Age	80.39	80.42	—	—
Illness burden score*	0.71	0.55	—	—
Inpatient care	$3734	$2072	$981	<.001
Outpatient care	$2920	$1983	$288	<.001
Outpatient pharmacy	$3557	$1443	$1902	<.001
Total	$10,212	$5498	$3171	<.001
% of entire cohort 2004 cost	36.64%	26.51%	—	—
Top 20% of entire cohort				
Age	79.79	80.26	—	—
Illness burden score*	2.32	2.11	—	—
Inpatient care	$8111	$8663	$308	0.545
Outpatient care	$8836	$9043	−$1081	<.001
Outpatient pharmacy	$4873	$3447	$1412	<.001
Total	$21,819	$21,153	$639	0.285
% of entire cohort 2004 cost	31.32%	40.80%	—	—
Top 5% of entire cohort				
Age	79.54	80.17	—	—
Illness burden score*	3.22	3.10	—	—
Inpatient care	$10,682	$12,649	$480	0.720
Outpatient care	$13,039	$14,646	−$2540	0.001
Outpatient pharmacy	$4822	$3645	$1172	<0.001
Total	$28,544	$30,941	−$888	0.575
% of entire cohort 2004 cost	10.25%	14.92%	—	—
Ratio of top 5% mean costs to bottom 50%	2.80	5.63		

*Based on DCG Model with AD indicators zeroed out and normalized to the full population.

†The risk-adjusted difference is calculated as $\beta 1$ from a weighted least square regression of the form $Y = \alpha + \beta * (\text{AD Cohort}) + \gamma_1 * (\text{RRS}) + \gamma_2 * (\text{RRS}^2)$, where Y is total cost, AD indicates being in the AD Cohort and RRS is the illness burden score.

‡The t test on the estimated $\beta 1$ coefficient, in the regression described below, for risk-adjusted cost.

them to costs for the bottom 50%. As seen in Table 3, the mean cost for the top 5% of the AD cohort is $28,544, accounting for 10% of all AD cohort cost. The top 5% of the control cohort has similar mean cost ($30,941), but it accounts for nearly 15% of the full control groups' overall cost. The bottom 50% of the AD cohort has mean annualized cost of $10,212, accounting for 37% of total AD cohort cost; the analogous figures for the controls are $5498 and 27%. Mean cost in the top 5% of the AD cohort is 2.8 times as high as for its least sick 50%; among the controls, this ratio is 5.6.

Average illness burden also varies dramatically within each cohort between the sickest 5% and bottom 50%. Mean RRS is 3.22 for the sickest 5% of AD patients and 0.71 for the bottom 50%, with the top group thus being 4.5 "times as sick"; the analogous figures among the controls are 3.10, 0.55, and 5.6.

Prevalence

The AD cohort has more non-AD illness than the controls, even though the cohorts are similar with respect to

TABLE 4. Prevalence and 2004 Mean Cost for AD and Control Cohorts by 2003 DCG Aggregated Condition Category (ACC)

	Prevalence			AD Cohort (Obs. = 25,109)		Control Group (Obs. = 75,327)		Excess AD Cost	
ACC	AD (%)	Control (%)	Odds Ratio†	Mean Actual Cost	Mean RRS	Mean Actual Cost	Mean RRS	Raw Difference	Risk-Adjusted Difference
				$13,936	1.23	$10,369	1.04	$3567*	$2307*
Infectious and parasitic	19.6	11.3	1.91*	$16,180	1.58	$15,047	1.54	$1133	$1587*
Malignant neoplasm	9.8	11.5	0.84*	$18,214	1.91	$17,358	1.79	$855	$495
Benign/in situ/uncertain neoplasm	12.9	15.9	0.78*	$15,690	1.36	$11,879	1.16	$3812*	$2453*
Diabetes	14.3	12.1	1.21*	$19,607	1.84	$16,081	1.64	$3526*	$2522*
Nutritional and metabolic	26.4	21.9	1.28*	$15,981	1.52	$12,954	1.32	$3027*	$2033*
Gastrointestinal	1.9	1.9	1.24*	$16,977	1.65	$14,654	1.49	$2324*	$1822*
Musculoskeletal and connective tissue	25.5	21.7	1.20*	$16,228	1.50	$13,090	1.35	$3139*	$2214*
Hematological	45.8	41.4	1.38*	$20,840	2.12	$22,268	2.14	−$1428	$1128
Cognitive disorders	11.6	8.7	155.72*	$14,289	1.31	$14,866	1.92	−$578	$3840*
Mental	53.1	0.7	5.08*	$17,396	1.68	$14,651	1.64	$2745*	$2448*
Neurological	0.6	0.3	2.32*	$18,393	1.76	$15,252	1.63	$3142*	$2740*
Cardio-respiratory arrest	15.4	3.5	1.43*	$22,604	2.42	$23,855	2.39	−$1251	$2105
Heart	0.1	0.0	1.16*	$15,815	1.48	$13,356	1.31	$2460*	$1842*
Cerebrovascular	14.7	6.9	2.60*	$17,481	1.69	$15,544	1.62	$1937*	$2129
Vascular	2.4	1.7	1.46*	$18,111	1.81	$16,988	1.76	$1123	$1261*
Lung	59.4	55.9	1.30*	$17,746	1.78	$16,408	1.66	$1338*	$1445*
Eyes	19.7	8.6	0.82*	$15,758	1.43	$12,076	1.22	$3682*	$2318*
Ears, nose, throat, and dental	17.1	12.4	0.99	$16,150	1.48	$13,116	1.30	$3034*	$2158*
Urinary system	24.8	20.3	1.59*	$18,824	1.81	$17,570	1.73	$1254	$1604*
Genital system	38.8	43.6	0.90*	$16,794	1.51	$12,617	1.28	$4177*	$2434*
Skin and subcutaneous	20.4	20.5	1.12*	$15,949	1.51	$13,087	1.32	$2861*	$1936*
Injury, poisoning, and complications	21.6	14.8	1.79*	$16,833	1.62	$14,720	1.53	$2112*	$2143*
Symptoms, signs, and ill-defined conditions	10.7	11.8	1.87*	$15,721	1.47	$13,930	1.40	$1791*	$1838*
Transplants, openings, and other V-codes	0.0	0.0	1.44*	$18,372	2.44	$20,833	2.33	−$2462	−$1647
Screening/history	31.1	28.7	0.83*	$16,354	1.49	$12,645	1.30	$3709*	$2378*

*Statistically significant at the 5% level: χ^2 test for prevalence; 2-sample t test for raw difference; t test for regression coefficient for risk-adjusted difference.
†Odds ratio = the odds for an AD person having this condition divided by the odds for a control having it.

age, sex, region, and eligibility. On average, those in the AD cohort have diseases across 8.1 nondementia CCs versus 6.5 for controls (Table 1).

We used odds ratios to quantify differences in prevalence rates for specific diseases between the AD and control cohorts. Odds ratios were computed by dividing the odds of the disease in the AD cohort by the odds of the disease in the control group. In general, an odds ratio greater than 1 means that the disease is more common among people with AD, whereas an odds ratio less than 1, means that it is less common for people with AD. Most ACCs are more common in the AD cohort (Table 4). Excess risk for AD patients is greatest (with odds ratios greater than 2) for the following conditions: cognitive disorders, mental health conditions, neurologic disorders, and cerebrovascular disease. However, some ACCs are less common for AD patients: malignant neoplasm, benign/in situ/uncertain neoplasm, eye conditions, genital system disorders, and screening/history.

At the more granular level, we report differences in CC prevalence and risk-adjusted costs for some particularly interesting comorbid conditions. All reported CCs are at least 1% prevalent in both the AD and control cohorts (Table 5). The selected CCs have either exceptionally different prevalence in the 2 cohorts, have high risk-adjusted differences in cost, or are particularly common.

The 5 most prevalent CCs in the AD cohort are major depressive, bipolar, and delusional disorders, depression, ischemic, or unspecified stroke, miscellaneous psychiatric disorders, and hip fracture or dislocation. AD patients have 5 times more risk of having major depressive, bipolar, and delusional disorders than the controls. We will report later that this mental condition is also the top in excess cost attributable to AD patients. There are a few CCs that are significantly less common in the AD cohort. These include: pelvic inflammatory disease/other specified female genital disorders, benign neoplasms of skin, breast and eye, glaucoma, history of disease, and lymphatic/head/neck/brain/major cancer. Within the highly common comorbid conditions, AD patients have more diabetes with no/unspecified complications, other endocrine/metabolic/nutritional disorders, miscellaneous gastrointestinal disorders, specified heart arrhythmias, precerebral arterial occlusion and transient cerebral ischemia, miscellaneous circulatory disease, viral and unspecified pneumonia: pleurisy, urinary tract infection, and miscellaneous injuries. AD patients, however, are less likely to have osteoarthritis of hip or knee and screening/observation exams.

TABLE 5. The Top Prevalence and Common Comorbidities for AD and Control Cohorts by 2003 DCG Condition Category (CC)*

	Prevalence			AD		Control		
CC	AD	Control	Odds Ratio‡	Mean Actual Cost	Mean RRS	Mean Actual Cost	Mean RRS	Risk-Adjusted Cost Diff.
Top 5 excess prevalence in ad cohort								
Major depressive, bipolar, and delusional disorders	5.0%	1.0%	5.1†	$18,616	1.9	$15,324	1.8	$3455†
Depression	4.8%	1.1%	4.5†	$17,807	1.6	$14,889	1.6	$1543
Ischemic or unspecified stroke	8.4%	2.6%	3.4†	$18,065	1.8	$16,894	1.7	$1354
Miscellaneous psychiatric disorders	3.5%	1.1%	3.2†	$15,475	1.7	$14,623	1.6	$755
Hip fracture/dislocation	4.2%	1.5%	2.8†	$16,508	1.7	$17,299	1.8	−$59
Top 5 excess prevalence in control cohort								
Pelvic inflammatory disease and other specified female genital disorders	1.1%	1.6%	0.7†	$15,900	1.4	$12,376	1.3	$3326†
Benign neoplasms of skin, breast and eye	8.0%	10.6%	0.7†	$16,920	1.4	$12,560	1.2	$2831†
Glaucoma	9.1%	11.8%	0.7†	$15,807	1.4	$11,912	1.2	$2241†
History of disease	5.0%	6.5%	0.8†	$16,226	1.5	$13,166	1.4	$2519†
Lymphatic, head, neck, brain and major cancer	1.2%	1.6%	0.8†	$21,463	2.4	$21,183	2.4	$57
Top 5 risk-adjusted differences in cost								
Polyneuropathy	1.9%	1.5%	1.2†	$24,033	2.1	$17,213	1.8	$4758†
Diabetes with neurological or peripheral circulatory manifestation	2.0%	1.5%	1.4†	$24,867	2.2	$19,060	2.1	$4476†
Major depressive, bipolar, and delusional disorders	5.0%	1.0%	5.1†	$18,616	1.9	$15,324	1.8	$3455†
Mononeuropathy, other neurological conditions/injuries	4.4%	3.8%	1.2†	$18,472	1.7	$13,665	1.5	$3389†
Pelvic inflammatory disease and other specified female genital disorders	1.1%	1.6%	0.7†	$15,900	1.4	$12,376	1.3	$3326†
Other selected common conditions								
Diabetes with no/unspecified complications	13.3%	11.3%	1.2†	$19,465	1.8	$16,088	1.6	$2330†
Other endocrine, metabolic, and nutritional disorders	20.5%	19.3%	1.1†	$15,650	1.4	$12,012	1.2	$2592†
Miscellaneous gastrointestinal disorders	23.1%	20.0%	1.2†	$17,040	1.6	$14,497	1.5	$2000†
Osteoarthritis of hip or knee	5.8%	6.9%	0.8†	$19,355	1.7	$14,417	1.5	$2904†
Precerebral arterial occlusion and transient cerebral ischemia	9.6%	5.8%	1.7†	$18,520	1.7	$15,708	1.6	$3047†
Miscellaneous circulatory disease	8.0%	5.5%	1.5†	$19,199	1.9	$16,696	1.8	$2038†
Viral and unspecified pneumonia and pleurisy	8.6%	4.8%	1.9†	$18,976	2.0	$19,209	2.0	$1730†
Urinary tract infection	11.0%	6.2%	1.9†	$17,210	1.7	$15,246	1.5	$1465†
Miscellaneous injuries	22.9%	14.2%	1.8†	$16,767	1.6	$14,064	1.5	$2597†
Screening/observation/dpecial exams	22.5%	27.7%	0.8†	$15,561	1.4	$11,789	1.2	$2746†

*At least 1% of prevalence rate in both the AD and control cohorts.
†Statistically significant at the 5% level.
‡Odds ratio = the odds for an AD person having this condition divided by the odds for a control having it.

Excess Costs Attributable to AD Patients

As noted in Table 2, costs for the AD cohort are 34% higher than for the control group. A key question is the extent to which higher non-AD morbidity accounts for the difference in cost. As described in the Methods section, we used a WLS model to estimate the incremental cost of AD. Estimated excess costs are $2307 per year, with pharmacy costs contributing $1711 to the total. We also calculated excess costs by type of service within 3 cost groups (the bottom 50%, top 20%, and top 5%) in Table 3. In general, excess costs attributable to AD decrease from the bottom (least sick) group to the top. Pharmacy cost is the key driver in the bottom group, accounting for more than 50% of the excess total.

We then examined differences in risk-adjusted costs by ACCs and CCs. For each clinical grouping, we computed a "risk-adjusted difference" in cost (that is, an estimated excess cost attributable to AD) and calculated its statistical significance (Table 4). The following ACCs have the highest excess costs for AD patients: cognitive disorder, neurologic disorder, diabetes, benign/in situ/uncertain neoplasm, and mental health disease. For example, the control group with neurologic disorder is sicker than the AD cohort (prospective RRS of 1.76 vs. 1.63). On the other hand, AD patients had a higher likelihood of having this condition recorded on their medical claims (odds ratio: 2.32). After adjusting for the differences in non-AD illness burden and the prevalence rate between the 2 groups, the estimated extra cost for AD patients with a neurologic disorder is both large ($2740) and significant.

Analyses at the CC level show similar patterns (Table 5). The 5 CCs with the highest excess costs for AD cohort are polyneuropathy, diabetes with neurologic/peripheral circulatory manifestation conditions, major depressive, bipolar, and

delusional disorders, mononeuropathy/other neurologic condition/injuries, and pelvic inflammatory disease, and other specified female genital disorders. The risk-adjusted cost attributable to AD patients with the most prevalent condition, major depressive, bipolar, and delusional disorders, is much higher ($3455) than the average excess total cost ($2307). Within the selected common comorbid conditions, excess costs attributable to AD patients ranged from $1465 to $3047.

DISCUSSION

In our study, AD patients are both sicker with non-AD conditions and more expensive than demographically matched controls. Even after adjusting for differences in non-AD illness burden, healthcare costs are higher for AD patients, with excess total costs attributable to AD averaging $2307 per year. Most of the higher cost is attributable to drug spending, particularly among AD patients with lower overall non-AD illness burden, although drug therapies might reduce some of the economic burden to the extent that they are able to delay the onset of other excess costs associated with the disease by delaying progression to more severe stages. We could not examine this with our cross sectional data. With the introduction of the Medicare Part D program in 2006, patterns of pharmacy use are likely to change for both AD and non-AD patients,[32] and program providers may attempt to "cherry-pick," that is, to avoid high-cost AD patients with significant comorbidities (ie, people who will be unprofitable because they are likely to need costly services). Additional research to answer such questions may assist the Center for Medicare and Medicaid Services (CMS) in designing more effective strategies for ensuring appropriate coverage and care for patients with AD.

The link between depression/depressive symptoms and the development of cognitive decline or dementia is still under investigation. Some epidemiological studies suggest that depression or depressive symptoms may be a risk factor,[33–36] whereas others have failed to confirm these findings.[33,34,37–39] Regardless of the direction of causality, AD patients in our study are at 5 times the risk of having major depressive, bipolar, and delusional disorders than controls. Furthermore, AD patients with these depressive symptoms cost $3455 more than controls, even after controlling for differences in total illness burden. The AD cohort has more recorded comorbidity (especially, mental health conditions, neurologic disorders, and cerebrovascular disease), but less malignant neoplasm, benign/in situ/uncertain neoplasms, and eye conditions. We suspect that this might be an artifact of under-detection, because AD patients have less screening than controls. Or, AD could be under-diagnosed among people being treated for other conditions, such as cancers.

How might a diagnosis of AD affect the cost of managing other diseases? Clinical studies have established a causal link between dementia and several other diagnoses, such as hip fracture, injury, pneumonia, and stroke. More generally, dementia is a major determinant of future functional impairment.[40] People with AD may become disoriented, increasing their risk of falls, hip fracture, and serious head injuries. AD patients may also have difficulty swallowing food and liquids, which may cause inhalation of some food and drink, leading to pneumonia. The positive link between AD and stroke may be related to undertreatment with nonsteroidal antiinflammatory drugs (NSAIDs) owing to concerns over serious side-effects, including ulcers, stomach irritation, and life-threatening stomach bleeding.[41–44] Thus, AD may both contribute to the development of other diseases as well as complicate their care.

Several study limitations are attributable to the nature of the data. The study population is Medicare-eligible retirees with employer-sponsored supplemental coverage, mostly from large firms whose former employees are not necessarily nationally representative. Importantly, the study results underestimate the total cost of AD because most costs associated with skilled nursing, home health, or hospice care are not captured. Another limitation is that AD can only be definitively confirmed by postmortem pathologic analysis. However, when the ICD9 331.0 code (Alzheimer disease) is recorded, it is because a clinician believes that AD is present. To address the problem that AD may also be present but unrecognized we excluded patients with other dementia diagnoses, but not AD, from both the AD and control groups. Finally, our prevalence analyses relied on coded diagnoses, and lower prevalence of certain comorbid medical conditions in AD may be at least partly due to less screening and, therefore, less recognized disease.

Our study contributes to the literature on cost of AD by investigating the prevalence and consequences of differences in comorbidities for an AD- and nondemented control cohort. We also quantified the expected effect of other morbidities on cost, enabling an estimate of the additional costs attributable to AD itself. Our approach could also be used to specify capitation rates that better align provider incentives with each AD patient's total illness burden.[45]

ACKNOWLEDGMENTS

The authors appreciate the contributions of David Paquette, RN and Timur Turkdogan, MS.

REFERENCES

1. Evans DA, Funkenstein HH, Albert MS, et al. Prevalence of Alzheimer's disease in a community population of older persons: higher than previously reported. *JAMA*. 1989;262:2552–2556.
2. Hebert LE, Scherr PA, Bienias JL, et al. Alzheimer disease in the U.S. population: prevalence estimates using the 2000 census. *Arch Neurol*. 2003;60:1119–1122.
3. Brokmeyer R, Jonhson E, Ziegler-Graham K, et al. *Forecasting the Global Burden of Alzheimer Disease.* Johns Hopkins University, Dept. of Biostatistics Working Papers. Working Paper 130; 2007.
4. Statistics about Alzheimer's Disease. Available at: http://www.alz.org/alzheimers_disease_alzheimer_statistics.asp. Alzheimer's Association, 2007. Accessed January 15, 2007.
5. *Saving Lives, Saving Money: Dividends for Americans Investing in Alzheimer Research.* A report from the Lewin Group, commissioned by the Alzheimer's Association. Washington, DC; 2004.
6. Hay JW, Ernst RL. The economic costs of Alzheimer's disease. *Am J Public Health*. 1987;77:1169–1175.
7. Ernst RL, Hay JW. The US economic and social costs of Alzheimer's disease revisited. *Am J Public Health*. 1994;84:1261–1264.
8. Weiner M, Powe NR, Weller WE, et al. Alzheimer's disease under managed care: implications from Medicare utilization and expenditure patterns. *J Am Geriatr Soc*. 1998;46:762–770.

9. Leibson C, Owens T, O'Brien P, et al. Use of physician and acute care services by persons with and without Alzheimer's disease: a population-based comparison. *J Am Geriatr Soc*. 1999;47:864–869.
10. Menzin J, Lang K, Friedman M, et al. The economic cost of Alzheimer's disease and related dementias to the California Medicaid Program (Medi-Cal) in 1995. *Am J Geriatr Psychiatry*. 1999;7:300–308.
11. Martin BC, Ricci JF, Kotzan JA, et al. The net cost of Alzheimer disease and related dementia: a population-based study of Georgia Medicaid recipients. *Alzheimer Dis Assoc Disord*. 2000;14:151–159.
12. Fillit H, Hill JW, Futterman R. Health care utilization and costs of Alzheimer's disease: the role of co-morbid conditions, disease stage, and pharmacology. *Fam Med*. 2002;34:528–535.
13. Hill JW, Futterman R, Duttagupta S, et al. Alzheimer's disease and related dementias increase costs of comorbidities in managed Medicare. *Neurology*. 2002;58:62–70.
14. Joyce A, Zhao Y, Bowman L, et al. Burden of illness among patients with Alzheimer's disease in a commercially-insured population. *Alzheims Dement*. 2007;3:204–210.
15. Gutterman EM, Markowitz JS, Lewis B, et al. Cost of Alzheimer's disease and related dementia in managed-Medicare. *J Am Geriatr Soc*. 1999;47:1065–1071.
16. Newcomer R, Clay T, Luxenberg JS, et al. Misclassification and selection bias when identifying Alzheimer's disease solely from Medicare claims records. *J Am Geriatr Soc*. 1999;47:215–219.
17. Richards KM, Shepherd MD, Crismon ML, et al. Medical services utilization and charge comparisons between elderly patients with and without Alzheimer's disease in a managed care organization. *Clin Ther*. 2000;22:775–791.
18. Doraiswamy PM, Leon J, Cummings JL, et al. Prevalence and impact of medical comorbidity in Alzheimer's disease. *J Gerontol A Bio Sci Med Sci*. 2002;57:M173–M177.
19. Sloan FA, Taylor DH Jr. Effect of Alzheimer disease on the cost of treating other diseases. *Alzheimer Dis Assoc Disord*. 2002;16:137–143.
20. 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;40:373–382.
21. Alzheimer's disease facts and figures 2007. Available at: http://www.alz.org/national/documents/report_alzfactsfigures2007.pdf. Alzheimer's Association, 2007. Accessed December 1, 2007.
22. American Medical Association. *International Classification of Diseases*, 9th rev. Clinical Modification (1996). Dover, DE: American Medical Association, 1995.
23. Academic Research License for DxCG's Diagnostic Cost Groups (DCGs) Models is available to obtain at http://www.dxcg.com/research-resources/academic-research-license.asp.
24. Ellis RP, Ash AS. Refinements to the Diagnostic Cost Group (DCG) Model. *Inquiry*. 1995;32:418–429.
25. Ellis RP, Pope GC, Iezzoni LI, et al. Diagnosis-based risk adjustment for Medicare capitation payments. *Health Care Financ Rev*. 1996;17:101–128.
26. Ash AS, Ellis RP, Pope GC, et al. Using diagnoses to describe populations and predict costs. *Health Care Financ Rev*. 2000;21:7–28.
27. Pope GC, Ellis RP, Ash AS, et al. Principal inpatient diagnostic cost group model for Medicare risk adjustment. *Health Care Financ Rev*. 2002;21:93–118.
28. Zhao Y, Ash AS, Ellis RP, et al. Disease burden profiles: an emerging tool for managing managed care. *Health Care Manag Sci*. 2002;5:211–219.
29. Pope GC, Kautter J, Ellis RP, et al. Risk adjustment of Medicare capitation payments using the CMS-HCC model. *Health Care Financ Rev*. 2004;25:119–141.
30. Zhao Y, Ash AS, Ellis RP, et al. Predicting pharmacy costs and other medical costs using diagnoses and drug claims. *Med Care*. 2005;43:34–43.
31. Winkelman R, Mehmud S. A comparative analysis of claims-based tools for health risk assessment. A research study sponsored by the Society of Actuaries (SOA). Available at http://www.soa.org/files/pdf/risk-assessmentc.pdf. Accessed April 20, 2007.
32. Introducing Medicare's new coverage for prescription drugs. Available at: http://www.medicare.gov/Publications/Pubs/pdf/11103.pdf. Centers for Medicare and Medicaid Services, 2005.
33. Speck CE, Kukull WA, Brenner DE, et al. History of depression as a risk factor for Alzheimer's disease. *Epidemiology*. 1995;6:366–369.
34. Green RC, Cupples LA, Kurz A, et al. Depression as a risk factor for Alzheimer Disease. *Arch Neurol*. 2003;60:753–759.
35. Steffens DC, Plassman BL, Helms MJ, et al. A twin study of late-onset depression and apolipoprotein E epsilon 4 as risk factors for Alzheimer's disease. *Biol Psychiatry*. 1997;41:851–856.
36. Ownby RL, Crocco E, Acevedo A, et al. Depression and risk for Alzheimer disease: systematic review, meta-analysis, and metaregression analysis. *Arch Gen Psychiatry*. 2006;63:530–538.
37. Ganguli M, Du Y, Dodge HH, et al. Depressive symptoms and cognitive decline in late life: a prospective epidemiological study. *Arch Gen Psychiatry*. 2006;63:153–160.
38. Lindsay J, Laurin D, Verreault R, et al. Risk factors for Alzheimer's disease: a prospective analysis from the Canadian Study of Health and Aging. *Am J Epidemiol*. 2002;156:445–453.
39. Chen P, Ganguli M, Mulsant BH, et al. The temporal relationship between depressive symptoms and dementia: a community-based prospective study. *Arch Gen Psychiatry*. 1999;56:261–266.
40. Agüero-Torres H, Fratiglioni L, Guo Z, et al. Dementia is a major cause of functional dependence in the elderly: 3 year followup data from a population-based study. *Am J Public Health*. 1998;88:1452–1456.
41. Finucane TE, Christmas C, Travis K. Tube feeding in patients with advanced dementia: a review of the evidence. *JAMA*. 1999;282:1365–1370.
42. Moroney JT, Tseng C-L, Paik MC, et al. Treatment for the secondary prevention of stroke in older patients: the influence of dementia status. *J Am Geriatr Soc*. 1999;47:824–829.
43. McGeer PL, Schulzer M, McGeer EG. Arthritis and anti-inflammatory agents as possible protective factors for Alzheimer's disease: a review of 17 epidemiologic studies. *Neurology*. 1996;47:425–432.
44. Aisen PS. The potential of anti-inflammatory drugs for the treatment of Alzheimer's disease. *Lancet Neurol*. 2002;1:279–284.
45. Newhouse JP, Buntin MB, Chapman JD. Risk adjustment and Medicare: Taking a close look. *Health Aff* (*Millwood*). 1997;16:26–43.

General Internal Medicine
Boston University School of Medicine
2008 Publications - L-Z

1. Levine SA, Chao SH, Brett B, **Jackson AH**, Burrows AB, Goldman LN, Caruso LB. Chief resident immersion training in the care of older adults: An innovative interspecialty education and leadership intervention. J Am Geriatr Soc. 2008;56(6):1140-45.
2. Levitzky YS, Cupples LA, **Murabito JM**, Kannel WB, Kiel DP, Wilson PWF, Wolf PA, O'Donnell CJ. Prediction of intermittent claudication, ischemic stroke, and other cardiovascular disease by detection of abdominal aortic calcific deposits by plain lumbar radiographs. Am J Cardiol. 2008;101(3):326-31.
3. Levitzky YS, Pencina MJ, D'Agostino RB, Meigs JB, **Murabito JM**, Vasan RS, Fox CS. Impact of impaired fasting glucose on cardiovascular disease: The framingham heart study. J Am Coll Cardiol. 2008;51(3):264-70.
4. **Liebschutz J**, **Battaglia T**, Finley E, **Averbuch T**. Disclosing intimate partner violence to health care clinicians - what a difference the setting makes: A qualitative study. BMC Public Health. 2008;8:229.
5. Lincoln A, Espejo D, Johnson P, **Paasche-Orlow M**, Speckman JL, Webber TL, White RF. Limited literacy and psychiatric disorders among users of an urban safety-net hospital's mental health outpatient clinic. J Nerv Ment Dis. 2008;196(9):687-93.
6. Lincoln A, Johnson P, Espejo D, Plachta-Elliott S, Lester P, **Shanahan C**, Abbott S, Cabral H, Jamanka A, Delman J, Kenny P. The BMC ACCESS project: The development of a medically enhanced safe haven shelter. J Behav Health Serv Res. 2008;36(4):478-91.
7. Lopes RD, Piccini JP, **Hylek EM**, Granger CB, Alexander JH. Antithrombotic therapy in atrial fibrillation: Guidelines translated for the clinician. J Thromb Thrombolysis. 2008;26(3):167-74.
8. **McCall-Hosenfeld JS**, **Freund KM**, Legault C, Jaramillo SA, Cochrane BB, Manson JE, Wenger NK, Eaton CB, McNeeley SG, Rodriguez BL, Bonds D. Sexual satisfaction and cardiovascular disease: The women's health initiative. Am J Med. 2008;121(4):295-301.
9. **McCall-Hosenfeld JS**, Jaramillo SA, Legault C, **Freund KM**, Cochrane BB, Manson JE, Wenger NK, Eaton CB, Rodriguez BL, McNeeley SG, Bonds D. Members of Women's Health Initiative-observational Study. Correlates of sexual satisfaction among sexually active postmenopausal women in the women's health initiative-observational study. J Gen Intern Med. 2008;23(12):2000-9.
10. McCarthy EP, Pencina MJ, Kelly-Hayes M, Evans JC, Oberacker EJ, D'Agostino RB S, Burns RB, **Murabito JM**. Advance care planning and health care preferences of community-dwelling elders: The Framingham Heart Study. J Gerontol A Biol Sci Med Sci. 2008;63(9):951-59.
11. Mehta SD, Hall J, **Greenwald JL**, Cranston K, Skolnik PR. Patient risks, outcomes, and costs of voluntary HIV testing at five testing sites within a medical center. Public Health Rep. 2008;123(5):608-17.
12. Merrick EL, Hodgkin D, Garnick DW, Horgan CM, Panas L, Ryan M, **Saitz R**, Blow FC. Unhealthy drinking patterns and receipt of preventive medical services by older adults. J Gen Intern Med. 2008;23(11):1741-48.

13. **Murabito JM**, Keyes MJ, Guo CY, Keaney JF, Jr, Vasan RS, D'Agostino RB S, Benjamin EJ. Cross-sectional relations of multiple inflammatory biomarkers to peripheral arterial disease: The framingham offspring study. Atherosclerosis. 2008;203(2):509-14.

14. **Murabito JM**, Pencina MJ, Zhu L, Kelly-Hayes M, Shrader P, D'Agostino RB S. Temporal trends in self-reported functional limitations and physical disability among the community-dwelling elderly population: The framingham heart study. Am J Public Health. 2008;98(7):1256-62.

15. **Paasche-Orlow MK**, Wolf MS. Evidence does not support clinical screening of literacy. J Gen Intern Med. 2008;23(1):100-102.

16. Peter I, Kelley-Hedgepeth A, Fox CS, Cupples LA, Huggins GS, Housman DE, Karas RH, Mendelsohn ME, Levy D, **Murabito JM**. Variation in estrogen-related genes associated with cardiovascular phenotypes and circulating estradiol, testosterone, and dehydroepiandrosterone sulfate levels. J Clin Endocrinol Metab. 2008;93(7):2779-85.

17. **Ramani S**. Twelve tips for excellent physical examination teaching. Med Teach. 2008;30(9-10):851-56.

18. **Ramani S**, Leinster S. AMEE guide no. 34: Teaching in the clinical environment. Med Teach. 2008;30(4):347-64.

19. **Rose AJ**, Ozonoff A, **Henault LE**, **Hylek EM**. Warfarin for atrial fibrillation in community-based practise. J Thromb Haemost. 2008;6(10):1647-54.

20. **Rose AJ**, Shimada SL, Rothendler JA, Reisman JI, Glassman PA, **Berlowitz DR**, **Kressin NR**. The accuracy of clinician perceptions of "usual" blood pressure control. J Gen Intern Med. 2008;23(2):180-83.

21. Rothman EF, **Cheng DM**, Pedley A, **Samet JH**, Palfai T, **Liebschutz JM**, **Saitz R**. Interpersonal violence exposure and alcohol treatment utilization among medical inpatients with alcohol dependence. J Subst Abuse Treat. 2008;34(4):464-70.

22. Rothman EF, DeJong W, Palfai T, **Saitz R**. Relationship of age of first drink to alcohol-related consequences among college students with unhealthy alcohol use. Subst Abus. 2008;29(1):33-41.

23. **Saitz R**. Poor care, not poor protocols, for alcohol withdrawal. Mayo Clin Proc. 2008;83(6):725-6; author reply 728-30.

24. **Saitz R**, Larson MJ, **LaBelle C**, Richardson JM, **Samet JH**. The case for chronic disease management for addiction. J Addic Med. 2008; 2:55-65.

25. **Saitz R**, **Horton NJ**, **Cheng DM**, **Samet JH**. Alcohol counseling reflects higher quality of primary care. J Gen Intern Med. 2008;23(9):1482-86.

26. **Samet JH**, Krupitsky EM, **Cheng DM**, **Raj A**, Egorova VY, Levenson S, **Meli S**, **Bridden C**, Verbitskaya EV, Kamb ML, Zvartau EE. Mitigating risky sexual behaviors among russian narcology hospital patients: The PREVENT (partnership to reduce the epidemic via engagement in narcology treatment) randomized controlled trial. Addiction. 2008;103(9):1474-83.

27. **Samet JH**, **Walley AY**. Can one be an expert in addiction medicine without expertise in pain management? Addiction. 2008;103(12):2006-7.

28. Seaver MR, **Freund KM**, Wright LM, Tjia J, Frayne SM. Healthcare preferences among lesbians: A focus group analysis. J Womens Health (Larchmt). 2008;17(2):215-25.

29. Siegel M, Albers AB, **Cheng DM**, Hamilton WL, Biener L. Local restaurant smoking regulations and the adolescent smoking initiation process: Results of a multilevel contextual analysis among massachusetts youth. Arch Pediatr Adolesc Med. 2008;162(5):477-83.

30. **Stillman MD**. Physicians behaving badly. JAMA. 2008;300(1):21-22.

31. Sullivan LE, **Saitz R**, **Cheng DM**, Libman H, Nunes D, **Samet JH**. The impact of alcohol use on depressive symptoms in human immunodeficiency virus-infected patients. Addiction. 2008;103(9):1461-67.

32. **Tsui JI**, Currie S, Shen H, Bini EJ, Brau N, Wright TL, VA HCV-001 Study Group. Treatment eligibility and outcomes in elderly patients with chronic hepatitis C: Results from the VA HCV-001 study. Dig Dis Sci. 2008;53(3):809-14.

33. Velagaleti RS, Pencina MJ, **Murabito JM**, Wang TJ, Parikh NI, D'Agostino RB, Levy D, Kannel WB, Vasan RS. Long-term trends in the incidence of heart failure after myocardial infarction. Circulation. 2008;118(20):2057-62.

34. Volandes AE, Ariza M, Abbo ED, **Paasche-Orlow M**. Overcoming educational barriers for advance care planning in latinos with video images. J Palliat Med. 2008;11(5):700-6.

35. Volandes AE, **Paasche-Orlow M**, Gillick MR, Cook EF, Shaykevich S, Abbo ED, Lehmann L. Health literacy not race predicts end-of-life care preferences. J Palliat Med. 2008;11(5):754-62.

36. Waite KR, **Paasche-Orlow M**, Rintamaki LS, Davis TC, Wolf MS. Literacy, social stigma, and HIV medication adherence. J Gen Intern Med. 2008;23(9):1367-72.

37. **Walley AY**, Alperen JK, **Cheng DM**, Botticelli M, Castro-Donlan C, **Samet JH**, **Alford DP**. Office-based management of opioid dependence with buprenorphine: Clinical practices and barriers. J Gen Intern Med. 2008;23(9):1393-98.

38. **Walley AY**, **Cheng DM**, Libman H, Nunes D, Horsburgh CR, Jr, **Saitz R**, **Samet JH**. Recent drug use, homelessness and increased short-term mortality in HIV-infected persons with alcohol problems. AIDS. 2008;22(3):415-20.

39. **Walley AY**, Krupitsky EM, **Cheng DM**, **Raj A**, Edwards EM, **Bridden C**, Egorova VY, Zvartau EE, Woody GE, **Samet JH**. Implications of cannabis use and heavy alcohol use on HIV drug risk behaviors in russian heroin users. AIDS Behav. 2008;12(4):662-69.

40. **Walley AY**, Phillips KA, Gordon AJ. The patients in recovery (PIR) perspective: Teaching physicians about methamphetamine. Subst Abus. 2008;29(4):61-64.

41. Wells KJ, **Battaglia TA**, Dudley DJ, Garcia R, Greene A, Calhoun E, Mandelblatt JS, Paskett ED, Raich PC, Patient Navigation Research Program. Patient navigation: State of the art or is it science? Cancer. 2008;113(8):1999-2010.

42. **Williams KN**, **Ramani S**, Fraser B, **Orlander JD**. Improving bedside teaching: Findings from a focus group study of learners. Acad Med. 2008;83(3):257-64.

43. Zhang Y, Guo X, **Saitz R**, Levy D, Sartini E, Niu J, Ellison RC. Secular trends in alcohol consumption over 50 years: The framingham study. Am J Med. 2008;121(8):695-701.

44. Zhao Y, Kuo TC, Weir S, Kramer MS, **Ash AS**. Healthcare costs and utilization for medicare beneficiaries with alzheimer's. BMC Health Serv Res. 2008;8:108.

Chief Resident Immersion Training in the Care of Older Adults: An Innovative Interspecialty Education and Leadership Intervention

Sharon A. Levine, MD,[*†‡] *Serena H. Chao, MD, MSc,*[*‡] *Belle Brett, EdD,*[§] *Angela H. Jackson, MD,*[*‖] *Adam B. Burrows, MD,*[*‡#] *Laura N. Goldman, MD,*[***] *and Lisa B. Caruso, MD, MPH*[*‡]

Chief residents (CRs) play a crucial role in training residents and students but may have limited geriatrics training or formal preparation for their CR role. A 2-day off-site chief resident immersion training (CRIT) addressed these challenges. Objectives were to foster collaboration between disciplines in the management of complex older patients, increase knowledge of geriatrics principles to incorporate into teaching, enhance leadership skills, and help CRs develop an achievable project for implementation in their CR year. Three cohorts totaling 47 trainees and 18 faculty mentors from 13 medical and surgical disciplines participated over 3 successive years. The curriculum, developed and taught by a multidisciplinary team, featured an interactive surgical case, mini-lectures on geriatrics topics, seminars to enhance teaching and leadership skills, and one-on-one mentoring to develop a project in geriatric care or education. Evaluation included pre- and postprogram tests and self-report surveys and two follow-up surveys or interviews. In 2006 and 2007, scores on a 12-item objective knowledge test increased significantly ($P<.001$) from before to immediately after CRIT. Self-report knowledge and confidence in teaching geriatrics also increased significantly ($P<.05$) in all formally covered topics. Mean enhancement of CR skills was 4.3 (1 = not at all, 5 = very much). Eleven months after CRIT, all but five CRs had implemented at least part of their action projects. CRs reported improved care of older patients, better leadership skills, more and better geriatrics teaching, and more collaboration between disciplines. A 2-day interactive program for CRs can increase institutional capacity regarding geriatrics teaching and care of elderly patients across medical specialties. **J Am Geriatr Soc 56:1140–1145, 2008.**

Key words: geriatrics education; internship and residency; graduate medical education; interdisciplinary communication; leadership

Chief residents (CRs) are uniquely positioned to influence peers and transform care. Senior faculty select CRs for their clinical, teaching, leadership, and interpersonal skills; trainees typically regard them with respect and affection.

Although CRs have had ample opportunities and guidance in the development of specialty-specific skills, their understanding of geriatrics issues may be limited.[1] They may not recognize the importance of physical and cognitive function in the care of frail older patients or be familiar with geriatric syndromes that pose risks in hospital settings. Furthermore, they may not have had opportunities to develop positive attitudes about the rewards of caring for this diverse, complex, and growing population.

With the well-documented projected increase of Americans aged 65 and older,[2] greater awareness and knowledge of geriatrics issues are vital for physicians in most specialties.[3] Equally important is the need for close collaboration and communication across specialties.[4] The term "interspecialty" is used to describe this collaborative effort that is necessary for the care of older adults.

In addition, CRs may receive little formal training in teaching techniques and leadership. Yet successful CRs must rely on solid teaching, communication, and leadership skills to fulfill their responsibilities as "flag bearers" for their departments and to have a positive effect on trainees.

To better prepare CRs for their new roles, several specialties have developed annual CR training programs focusing on leadership and management skills. Published reports indicate the effectiveness of these efforts.[5–7] Programs have also been developed to augment residents' geriatrics training within their specialties[8] and to provide CRs with enhanced clinical knowledge.[9] This article describes a 2-day educational intervention for CRs from multiple disciplines that combines training in leadership and teaching

From the [*]Department of Medicine; [†]Office of the Dean, School of Medicine, Boston University, Boston, Massachusetts; [‡]Geriatrics Section; [‖]Section of General Internal Medicine; [**]Department of Family Medicine, Boston Medical Center, Boston, Massachusetts; [§]Brett Consulting Group, Somerville, Massachusetts; and [#]Upham's Corner Health Center, Boston, Massachusetts.

Address correspondence to Sharon A. Levine, MD, Office of the Dean, 715 Albany Street, L-103, Boston, MA 02118. E-mail: salevine@bu.edu

DOI: 10.1111/j.1532-5415.2008.01710.x

0002-8614/08/$15.00

and in geriatrics while emphasizing the significance of collaboration in the care of older adults.

In November 2003, the Donald W. Reynolds Foundation awarded a 4-year grant to the Geriatrics Section at Boston University School of Medicine to develop the Boston University Medical Center Comprehensive Geriatric Education project. Its funding was intended to improve the care of older adults by increasing training in geriatrics in the many disciplines of medicine where older adults are seen and at all levels of training from students to practicing physicians. As part of this project, the Chief Resident Immersion Training Program in the Care of Older Adults (CRIT) was conceived as an educational program for CRs. Through CRIT, the innovative step of enlisting CRs to improve the teaching and dissemination of geriatrics knowledge within their departments has been taken.

The overall goal of CRIT was to foster interspecialty collaboration in the management of complex older patients. The specific objectives were to help CRs at Boston Medical Center (BMC) to incorporate geriatrics principles into their teaching and administrative roles, develop leadership and teaching skills for the care of complex older patients, enhance leadership and teaching skills required in the CR role, develop an achievable project to be conducted during the chief residency year, and have fun and foster collegiality.

METHODS

Development of the Curriculum

A multidisciplinary team of key faculty from the sections of Geriatrics and General Internal Medicine and the Department of Family Medicine at BMC met regularly to develop a curriculum based on a needs assessment of rising CRs as well as geriatrics content and leadership and teaching skills deemed important by the faculty.

As a central focus of the program, a fictional case was created of an 84-year-old community-dwelling woman with multiple medical conditions who presented to the emergency department with bowel perforation due to diverticulitis, underwent surgery, developed postoperative complications, and was then discharged. The case incorporated several key aspects of geriatric care: assessment of decision-making capacity, preoperative evaluation, recognition and management of perioperative issues including delirium, functional assessment, polypharmacy (added in 2006), and discharge planning. The case was composed so that details would be revealed over three 2-hour interactive modules designed to stimulate active learning of core concepts in geriatric care and to enhance teaching and leadership skills. Each module included small-group discussions, evidence-based mini-lectures on topics in geriatrics, and interactive seminars focused on CR skills. A facilitator's guide containing questions to elicit key concepts was also developed for use during small-group discussions.

Recruitment

CRs were invited from programs whose residents care for older patients. Program directors were also invited with the hope that they, too, would benefit from the geriatrics educational content of the weekend and support each CR in carrying out a yearlong action project.

Program

The CRIT weekend, held at a pleasant off-site location approximately 2 hours from Boston, began Friday night and ended on Sunday after lunch. To set the stage on the opening evening, in Years 2 and 3, there were guest speakers who described an older relative's illness or injury, highlighting the strengths and weaknesses of the medical care and the practitioners' communications with the patient, family, and each other. The fictional case discussion began the next morning. Two modules were completed on Saturday, and the final module was completed on Sunday. Each of the three modules started with a brief presentation of the case. Participants then engaged in 30-minute small-group discussions about issues raised in the case, facilitated by geriatricians and gerontologically trained advance practice nursing faculty members. In keeping with the central theme of interspecialty care, small groups consisted of program directors and incoming CRs from different specialties; participants were assigned to different small groups for each successive module. The large group then reconvened to receive two 15-minute evidence-based mini-lectures on topics raised in the small-group discussion. At the conclusion of each module, a faculty member led a 20-minute interactive seminar designed to enhance CR teaching and leadership skills that were linked to the management of the unfolding case. Topics included basic "survival skills" of the CR: giving feedback, approaching the reluctant learner, conflict resolution, and teaching in small groups.

Action Projects

Each CR was asked to develop a project action plan that focused on the management of complex older patients specific to needs of their specialty and residency program. The project, to be completed within the CR year, could involve trainee education, improvement in clinical care, institutional change, administration, or patient education. The CRs met one-on-one with faculty at two regularly scheduled times during the weekend to develop and refine their action plans. Selected projects were shared with the entire group near the end of CRIT.

EVALUATION

Instruments and Analysis

At the time of CRIT, the evaluation collected qualitative and quantitative data through pre- and postknowledge tests and pre- and post–self-report surveys. (Program directors were also given a short postsurvey, but results are not included in this article.) The pretest and presurvey were administered to participants before CRIT. The multiple-choice knowledge test, which was revised over each of the 3 years, consisted of 12 items, with seven or eight items that were mini-cases. In the pre- and postsurveys, CRs were asked to assess their knowledge of 19 topics related to geriatrics and their confidence in teaching these topics. To enhance validity, the pre- and postsurveys intentionally contained items on five topics not formally covered in the program. In the presurvey, CRs also rated their confidence

in a variety of skills related to their roles as CRs, and in the postsurvey, they rated the extent to which the program enhanced their knowledge of these skills.

Six months after the program, CRs were invited to a follow-up discussion about progress on their action projects and their use of geriatrics content and CR skills learned at CRIT. Those who did not attend were contacted for an interview on the same topics. In 2006, at the 6-month follow-up interview, a question was added about the relationships they had made with other disciplines. At 11 to 12 months after the program, all CRs were contacted and asked to complete a Web-based survey or respond to a phone interview. Questions were quantitative and qualitative and focused on career plans and completion of action projects. (At the time of this article, data were collected only for the 2005 and 2006 cohorts.)

Quantitative data were analyzed using SPSS (SPSS Inc., Chicago, IL). Two-tailed *t*-tests were used to compare means on pre and post questions for the combined 2006 and 2007 data. (The 2005 form used only retrospective pre and post ratings, which were kept in the later forms for comparison purposes.) Qualitative data were analyzed for themes.

RESULTS

Participants

A total of 47 trainees participated over all 3 years, including 44 CRs, two fellows in a surgical subspecialty program, and one second-year resident scheduled to be a CR. Yearly totals were 12 in 2005, 16 in 2006, and 19 in 2007. Seven program directors or faculty mentors accompanied these trainees in 2005, six in 2006, and five in 2007. (One program director attended all three times, and three others attended twice.) Over all 3 years, medical disciplines represented by the trainees included internal medicine (n = 8), family medicine (n = 4), neurology (n = 5), psychiatry (n = 5), rehabilitation medicine (n = 3), anesthesiology (n = 6), cardiothoracic surgery (n = 2), otolaryngology (n = 6), ophthalmology (n = 3), general surgery (n = 1), urology (n = 2), and emergency medicine (n = 2). Three disciplines (urology, psychiatry, and emergency medicine) never sent a faculty mentor; one discipline (obstetrics and gynecology) sent one faculty mentor without having any trainees participate.

Pre–Post Geriatrics Knowledge Test

In 2005 (response rate 100%), CRs answered 6.7 of 10 questions correctly in the objective pre–knowledge test and 7.2 in the posttest. The change was not significant ($P > .05$). In 2006 (response rate 73%, mainly because of missing pre data), with changes in the test to make it more relevant and more difficult, mean correct responses out of 12 questions went from 5.8 (95% confidence interval (CI) = 4.9–6.7) on the pretest to 8.4 (95% CI = 7.6–9.2) on the posttest ($P = .001$). In 2007 (response rate of 100%), mean correct responses out of 12 questions went from 6.7 (95% CI = 6.0–7.3) to 8.9 (95% CI = 8.3–9.6) ($P < .001$). Questions with greatest gains varied according to year.

Self-Report Change in Geriatrics Knowledge and Confidence in Teaching About Geriatrics

Combining data for 2006 and 2007, there were significant differences ($P < .001$) between respondents' assessments of their knowledge before and at the end of the program on all topics formally covered in the program. Differences on topics not covered formally were not significant, except for "assessing and reducing the risk for falls" ($P < .04$). Table 1 shows pre and post ratings and *P*-values for combined 2006 and 2007 participants.

For all 3 years combined, results for pre–post confidence in teaching geriatrics topics to trainees also show significant differences ($P < .001$) for all items covered in the formal curriculum. Differences between mean scores reached significance ($P < .05$) for all but one of the uncovered topics.

Confidence Levels and Enhancement of Teaching and Leadership Skills for a CR

Table 2 indicates a spread of CR initial confidence levels related to skills and knowledge needed to be a CR. The overall mean for initial confidence in CR skills was 3.6 (95% CI = 3.5–3.8, missing = 6), with 1 = not at all and 5 = very much. On the postsurvey, participants were asked the degree to which the program had enhanced their knowledge in these areas. The overall mean level of enhancement was 4.3 (95% CI = 4.1–4.5, missing = 1).

Follow-Up: Impact on Clinical Work

In the 2005 and 2006 follow-up, participants named a variety of ways in which CRIT had made a difference in the ways they thought about or practiced medicine. Specific concepts mentioned most often included better medication management (n = 6) and more awareness regarding diagnosing delirium (n = 6). The most common theme was a heightened sensitivity to the unique needs of older patients, requiring greater attention to their multidisciplinary needs, their different physiology, their function, and the home environment to which they would be returning, as well as more patience and better communication from practitioners. CRs from 2006 noted that CRIT made it easier to connect with the other disciplines regarding patient issues (n = 6).

Follow-Up: Effect on Work as CR

All 11 active CRs from 2005 and the 15 respondents from 2006 provided ratings on the effect of CRIT on their overall ability to perform their work as a CR. The mean rating was 3.69 (95% CI = 3.41–3.97), with only one person rating the effect as less than a 3 on a 5-point scale (1 = none to 5 = a great deal). In response to an open-ended question asking for specific examples of how participants had used what they learned at CRIT in their roles as CRs, both cohorts named enhanced conflict resolution skills (n = 14); in 2005, trainees also mentioned learning new concepts related to the care of older adults to teach to residents and students (n = 6) and, in 2006, teaching and leadership skills (n = 9).

Table 1. True Pre and Post Self-Reported Geriatrics Knowledge, Combined 2006 and 2007

Topic	N	Pre	Post	*P*-Value
		Mean ± Standard Deviation		
Topic covered formally				
Assessing decision-making in elderly patients	31	3.29 ± 0.94	4.37 ± 0.58	<.001
Conducting a preoperative assessment of an elderly patient	30	3.07 ± 1.31	3.97 ± 0.67	<.001
Conducting a functional assessment of older patients	31	3.06 ± 1.09	4.13 ± 0.67	<.001
Recognizing dementia	31	3.87 ± 0.89	4.42 ± 0.56	<.001
Managing dementia	31	3.16 ± 1.19	4.06 ± 0.77	<.001
Recognizing delirium	31	4.06 ± 0.89	4.81 ± 0.40	<.001
Managing delirium	31	3.65 ± 1.08	4.55 ± 0.62	<.001
Reviewing medications for evidence of polypharmacy	31	3.23 ± 1.02	4.27 ± 0.60	<.001
Assessing the adequacy of the patient's social support and living arrangements	30	3.40 ± 1.22	4.17 ± 0.65	.001
Creating a postdischarge management plan	31	3.32 ± 1.10	4.16 ± 0.74	<.001
Understanding implications of different insurance coverage for older patients	31	1.98 ± 1.14	3.52 ± 0.81	<.001
Knowledge of long-term care services, including home care services	31	2.70 ± 1.18	4.07 ± 0.83	<.001
Recognizing the value of and facilitating the interdisciplinary, collaborative team process	31	3.74 ± 1.03	4.68 ± 0.54	<.001
Incorporating the principles of geriatric rehabilitation	31	2.81 ± 1.08	3.97 ± 0.71	<.001
Topic not covered formally				
Recognizing, evaluating, and treating urinary incontinence and voiding difficulties	29	3.24 ± 1.06	3.10 ± 1.01	.42
Assessing and reducing risk for falls	31	3.42 ± 1.09	3.77 ± 0.92	.04
Managing diabetes mellitus	30	3.47 ± 1.17	3.60 ± 1.30	.29
Managing coronary artery disease	31	3.35 ± 1.20	3.55 ± 1.34	.06
Taking into account cultural differences in making decisions regarding patient care plans	30	3.50 ± 1.33	3.63 ± 0.96	.51

Based on a scale of 1 (low) to 5 (high).

Follow-Up: Action Projects

Twenty of 27 responding trainees from 2005 and 2006 had completed at least half of their action projects at the 11-month follow-up (mean completion rate 60%, 95% CI = 46–75%), with those in nonsurgical specialties completing a significantly greater percentage of their projects than those in surgical specialties (mean completion rate 75% vs 41%, $P = .01$). Eight individuals accomplished 100% of their projects; five had not completed any part of their projects. Barriers to completion included time, need for official approval, recruitment of a sufficient number of subjects, resistance to change among hospital personnel, lack of published data on a topic, lack of faculty buy-in, and performance of project somewhere else. Table 3 contains a sample list of action projects.

Table 2. Initial Confidence in and Extent of Enhancement of Chief Resident (CR) Skills, Combined 2005, 2006, and 2007

Skills	Confidence in Skills Related to Being a CR		Extent to Which Chief Resident Immersion Training Enhanced CR Skills	
	N	Mean ± SD (95% CI)	N	Mean ± SD (95% CI)
Group facilitation skills	42	3.71 ± 0.84 (3.45–3.97)	46	4.11 ± 0.92 (3.83–4.38)
Feedback skills	43	3.81 ± 0.73 (3.59–4.04)	46	4.33 ± 0.79 (4.09–4.56)
Connecting with a reluctant learner	43	3.37 ± 0.85 (3.11–3.63)	46	4.43 ± 0.75 (4.21–4.66)
Incorporating geriatrics issues into formal and informal teaching	43	3.26 ± 0.90 (2.98–3.53)	46	4.41 ± 0.69 (4.21–4.62)
Teaching others clinical problem-solving skills related to care of older patients	42	3.33 ± 0.93 (3.04–3.62)	46	4.20 ± 0.62 (4.01–4.38)
Leading a team	43	4.05 ± 0.87 (3.78–4.31)	46	4.22 ± 0.92 (3.95–4.49)
Resolving conflicts within a multidisciplinary team	43	3.74 ± 0.82 (3.49–4.00)	46	4.35 ± 0.67 (4.15–4.55)
Managing the multiple responsibilities of position as CR	43	3.86 ± 0.80 (3.61–4.11)	46	4.35 ± 0.77 (4.12–4.58)
Total of all items	41	3.63 ± 0.55 (3.46–3.81)	46	4.30 ± 0.61 (4.12–4.48)

Based on a scale of 1 (not at all) to 5 (very much).
SD = standard deviation; CI = confidence interval.

Table 3. Sample List of Proposed Action Plans

Specialty	Setting/Audience	Action Plan Project
Anesthesiology	Weekly conference/residents	"Understanding the Complexities of Managing the Elderly Patient in the Perioperative Setting"/To develop lecture and online resources with focus on managing elderly patients
Cardiothoracic surgery	Residents	To identify different surgical options available for lung cancer
Emergency medicine	Residents and attendings	"Pre-Op Assessment of the Geriatric Patient in the ED"/To educate ED staff to assess geriatric patients before operative repair of orthopedic injury
Family medicine	Community health center	To train residents and interns to review patients' healthcare proxies and end-of-life wishes
Internal medicine	Internal medicine intranet Web site/interns and residents	To incorporate evidence-based geriatrics bibliography into Web-based references for internal medicine trainees
Internal medicine	Intern conference/medical interns	"Dementia and Delirium"/To increase intern competence in the diagnosis and management of patients with these diagnoses
Internal medicine	Intern conference/interns and residents	"End-of-Life Care in Medically Complicated Elders"/To identify elderly people with end-of-life care needs. Choose appropriate means for symptom management. Combine cultural, medical, and social aspects into care plan
Neurology	Didactic lecture/neurology residents and medical students	"Gait Assessment and Fall Risk in Elderly Patients with Neurological Disease"/To understand different gait problems associated with neurological diseases, falls risk, and preventive measures
Ophthalmology	Residents and faculty	"Communication to Primary Care Providers Regarding Ophthalmic Disease"/To provide primary care physicians with regular updates and communication regarding their patients' visual and ophthalmic health
Otolaryngology	Otolaryngology grand rounds/residents in otolaryngology	"Prevention, Diagnosis and Management of Postoperative Delirium in ENT Patients"/To increase awareness in house staff about prevention and treatment of common postsurgical problems
Otolaryngology	Medicine and geriatrics departments grand rounds; Internal medicine Web-based journal/general internists, geriatricians	"Dysphagia: Cross-Disciplinary Diagnosis and Practical Management"/To produce and present a comprehensive talk with practical information on this topic for other disciplines
Psychiatry	Psychiatry grand rounds/attendings, residents, students	"Interdisciplinary and Community Resources for Caregivers of Patients with Dementia"/To inform psychiatry trainees and faculty about community-based resources for caregivers
Psychiatry	Ward rounds teaching/medical student psychiatry clerks and psychiatry residents	"Case-Based Approaches to Teaching Drug Interaction Concepts in the Elderly"/To minimize adverse drug interactions in psychiatric patients
Surgery	Anesthesia and surgical house staff	To improve recognition and treatment of delirium in postoperative and surgical intensive care unit patients
Urology	Urology residents and faculty	"Counseling Geriatric Patients on PSA and Prostate Cancer"/To increase awareness and knowledge of prostate cancer diagnosis and its efforts on geriatric patients

ED = emergency department; ENT = ear/nose/throat; PSA = prostate-specific antigen.

DISCUSSION

Although published reports exist documenting the success of weekend workshops designed to provide CRs with added administrative and leadership skills in specialties such as family medicine, pediatrics, and psychiatry,[5–7] the authors were unaware of any educational training for CRs across specialties designed to meet the stated objectives. CRIT in the Care of Older Adults provided an effective forum for raising CR awareness of geriatric issues, increasing knowledge about specific geriatric syndromes and practices, enhancing skill sets and building confidence for work as a CR, and reinforcing the importance of interdisciplinary and interspecialty approaches to the care of older patients. It also provided an off-site opportunity for CRs, program directors, and geriatrics faculty to establish collegial connections integral to professional collaboration in the care of complex patients. Follow-up of the 2005 and 2006 cohorts at 6 and 11 months suggests that CRIT affected CRs' medical practices with older patients and their awareness of the importance of interdisciplinary and interspecialty collaboration in managing the care of these patients. CRIT provided CRs with new concepts in geriatrics to teach students and residents and skills in conflict resolution, teaching, and leadership. All but five of the CRs were successful in completing at least part of their action projects during their CR year. Those who attended CRIT believed that it would affect the ways they performed their future work as physicians. Additional comments by participants suggested that, overall, the weekend was appropriately designed and organized to meet their needs.

Several factors contributed to the success of CRIT. Holding the program at an off-site location allowed participants to focus fully on accomplishing their educational goals. The use of mixed methods sustained the learners' attention. CRs felt comfortable learning with the program directors and faculty mentors, whose presence added validation to the experience and provided role models for continuing education. Informal learning about a broader

scope of geriatrics topics, such as falls and incontinence, occurred during small-group sessions and even during mealtime conversations. Finally, the leadership and teaching portion of the curriculum was able to cover the diverse roles for CRs across disciplines; some spent much time teaching, and others were responsible for scheduling operative procedures. Even so, they had similar concerns about their ability to manage the challenges of their leadership roles, and they found common ground and camaraderie in discussing these issues in small groups.

A few potential limitations should be considered. First, the program may be difficult to replicate. BMC has a strong geriatrics program, giving the program a visible clinical presence. It could be argued that this type of program may lack efficacy in an institution with few geriatrics faculty, although only four members of the Geriatrics Section were used while also engaging "geriatrics friendly" faculty from the Section of General Internal Medicine and the Department of Family Medicine.

Second, there are the usual limitations of small sample sizes and self-report data, but to increase the strength of these data, a broad range of outcomes was considered; quantitative and qualitative data were collected at different time frames, and a variety of methods and instruments was used, including an objective quantitative knowledge test, surveys before and immediately after the program, and follow-up surveys and interviews at 6 and 11 months after the program. The action projects undertaken by CRs served as an observed assessment of application of knowledge. The qualitative data provided a fuller understanding of the ways in which the CRs learned what they used at CRIT.

Third, it was not possible to more rigorously document behavior change on the part of the CRs. The resources were not available for chart audits and review of resident evaluations of CR performance. Fourth, it was not possible to document that significant changes were sustained 1 year or more after the completion of the CR year given that most CRs leave the institution and are hard to track down.

In final summary, CRs are an untapped resource for changing geriatrics practice and education, especially for increasing interspecialty communication and collaboration between providers of health care for frail and complex older patients. CRs can be a source of cross-fertilization between departments at an institution and are eager learners who often become leaders at other institutions. By focusing on this important group, CRIT is an efficient and effective means of increasing institutional capacity for enhancing awareness, knowledge, and skills regarding high-quality collaborative care of elderly patients.

ACKNOWLEDGMENTS

The authors would like to thank Jeffrey N. Katz, MD, for his thoughtful review of an earlier draft of the manuscript. The CRIT Program evolved from conversations between the authors and many valued colleagues who generously shared ideas and insights on chief resident training. We particularly and gratefully acknowledge Daniel P. Alford, MD, MPH, and Jeffrey H. Samet, MD, MA, MPH, for formative conversations that helped give shape to the program assessed in this manuscript. Drs. Katz, Alford, and Samet received no compensation outside of regular salary for their contributions.

Conflict of Interest: The editor in chief has reviewed the conflict of interest checklist provided by the authors and has determined that none of the authors have any financial or any other kind of personal conflicts with this article. The CRIT Program was supported by a grant from the Donald W. Reynolds Foundation. Drs. Levine, Chao, Jackson, Burrows, Caruso, Goldman, and Brett all receive funding from the Donald W. Reynolds Foundation.

Author Contributions: Sharon Levine and Serena Chao made substantial contributions to the study conception and design of the article and contributed significantly to the acquisition of subjects and data, analysis and interpretation of data, and drafting and critical revision of the manuscript. Belle Brett made substantial contributions to the study conception and design of the article and contributed significantly in the acquisition of data, and analysis and interpretation of data, and to the drafting and critical revision of the manuscript. Angela Jackson, Adam Burrows, Laura Goldman, and Lisa Caruso made substantial contributions to the study conception and design of the article and contributed significantly to the drafting and critical revision of the manuscript.

Sponsor's Role: The Reynolds Foundation was not involved in the design or conduct of the study; in the collection, management, analysis, or interpretation of the data; or in the preparation, review, or approval of the manuscript.

REFERENCES

1. Solomon DH, Burton JR, Lundebjerg NE et al. The new frontier: Increasing geriatrics expertise in surgical and medical specialties. J Am Geriatr Soc 2000;48:702–704.
2. He W, Sangupta M, Velkoff VA et al. US Census Bureau, Current Population Reports, P23-209, *65+in the United States: 2005*, U.S. Government Printing Office, Washington, DC, 2005.
3. A statement of principles: Towards improved care of older patients in surgical and medical specialties. J Am Geriatr Soc 2000;48:699–701.
4. High KP, Clayton CP, Woolard N et al. Expanding geriatrics into subspecialty internal medicine healthcare, research, and education: Caring for older adults and the T. Franklin Williams career development awards. Am J Med 2002; 113:533–536.
5. Mygdal WK, Monteiro M, Hitchcock M et al. Outcomes of the first family practice chief resident leadership conference. Fam Med 1991;23:308–310.
6. Doughty RA, Williams PD, Seashore CN. Chief resident training: Developing leadership skills for future medical leaders. Am J Dis Children 1991;145: 639–642.
7. Wilder JF, Plutchik R, Conte HR. The role of the chief resident: Expectations and reality. Am J Psychiatry 1976;133:328–331.
8. Potter JF, Burton JR, Drach GW et al. Geriatrics for residents in the surgical and medical specialties: Implementation of curricula and training experiences. J Am Geriatr Soc 2005;54:511–515.
9. Alford DP, Clark TW, Samet JH. Promoting substance abuse education among generalists: The Chief Resident Immersion (CRIT) program. J Gen Interm Med 2004;19(Suppl 1):93.

Prediction of Intermittent Claudication, Ischemic Stroke, and Other Cardiovascular Disease by Detection of Abdominal Aortic Calcific Deposits by Plain Lumbar Radiographs

Yamini S. Levitzky, MD[a], L. Adrienne Cupples, PhD[b], Joanne M. Murabito, MD, ScM[a,d], William B. Kannel, MD[a], Douglas P. Kiel, MD, MPH[e], Peter W.F. Wilson, MD[g], Philip A. Wolf, MD[a,c], and Christopher J. O'Donnell, MD, MPH[a,f,*]

There has been little attention to vascular calcium testing for generalized assessment of cardiovascular disease (CVD) outcomes, such as intermittent claudication (IC) and ischemic stroke (IS). We hypothesize that aortic calcium is an important predictor of CVD outcomes. Lumbar x-rays were obtained in 848 men and 1,301 women (mean ages 59.7 and 60.1 years, respectively) from the original cohort of the Framingham Heart Study. Abdominal aortic calcium (AAC) deposits were graded using a previously validated scale. Participants were categorized according to a 10-year Framingham coronary heart disease (CHD) risk score. Multivariable Cox proportional hazards analyses were performed to relate AAC to CVD outcomes. There were 199 IC events, 201 IS events, 702 CHD events, and 1,121 CVD events during 32 years of follow-up. Multivariable adjusted hazard ratios for the third versus first AAC tertile in the combined cohort were 1.68 (95% confidence interval [CI] 1.12 to 2.50) for IC, 1.73 (95% CI 1.12 to 2.65) for IS, 1.59 (95% CI 1.26 to 2.00) for CHD, and 1.64 (95% CI 1.37 to 1.97) for CVD. Hazard ratios for IC and IS were similar in magnitude to those for CHD and CVD. A high AAC score was associated with significantly higher incidence of events in subjects at intermediate Framingham CHD risk for all end points. Risk prediction based on cardiovascular risk factors improved for most outcomes when AAC was added. In conclusion, there was a graded, increasing, and independent association of AAC with incident IC and IS, similar in magnitude to risks predicted for CHD and CVD. AAC appears to be useful for risk stratification in patients at intermediate CHD risk. **(Am J Cardiol 2008;101:326–331)**

One easily detected manifestation of subclinical atherosclerosis is vascular calcium. Atherosclerosis begins to develop in the aorta before it appears in the lower extremity, cerebral, or coronary vascular beds.[1] Abdominal aortic calcium (AAC)[2,3] and aortic arch calcium,[4,5] detected by plain radiography or computed tomography, are manifestations of aortic atherosclerotic plaque associated with increased cardiovascular disease (CVD) morbidity and mortality in prospective epidemiologic studies. Previous studies have documented an association between aortic arch calcium and risk of peripheral arterial disease.[4,5] Additionally, the Rotterdam Coronary Calcification Study demonstrated an association between presence of aortic calcium (detected by computed tomography) and coronary calcium.[6] We previously reported AAC is a predictor of coronary heart disease (CHD), CVD, and CVD mortality,[2] and similar findings were noted for CVD mortality in the Rotterdam study.[3] We took advantage of an existing study of lateral lumbar x-rays conducted in the original cohort of the Framingham Heart Study (FHS) in 1966 to 1970. In the present study, we hypothesized that AAC is a useful marker for systemic atherosclerosis in many vascular beds and therefore may be predictive of future non-CHD events, such as intermittent claudication (IC) and ischemic stroke (IS), and CHD and total CVD, over and above the traditional risk factors, in particular in patients at intermediate risk for CHD.

Methods

Subjects were participants of the FHS, a prospective cohort study initiated in 1948 designed to elucidate risk factors for CHD. The original cohort consisted of 5,209 volunteers from Framingham, Massachusetts, 28 to 62 years of age at enrollment who have continued to undergo biennial exam-

[a]National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, and [b]Department of Biostatistics, Boston University School of Public Health, [c]Departments of Neurology, Preventive Medicine and Epidemiology and [d]Section of General Internal Medicine, Boston University School of Medicine, and [e]Institute for Aging Research Hebrew SeniorLife and [f]Cardiology Division, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts; and [g]Division of Cardiology, Emory University School of Medicine, Atlanta, Georgia. Manuscript received April 1, 2007; revised manuscript received and accepted August 20, 2007.

This work was supported by Grant N01-HC-38038 from the National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, Massachusetts; Precursors of Stroke Incidence and Prognosis Grant NIH/NINDS 2R01 NS017950 from the National Institute for Neurological Disorders and Stroke, Bethesda, Maryland; and Grant R01 AR/AG 41398 from the National Institute of Arthritis, Musculoskeletal and Skin Diseases and National Institute on Aging, Bethesda, Maryland.

*Corresponding author: Tel: 508-935-3435; fax: 508-626-1262.
E-mail address: odonnellc@nhlbi.nih.gov (C.J. O'Donnell).

0002-9149/08/$ – see front matter
doi:10.1016/j.amjcard.2007.08.032

inations. The study design and inclusion criteria have been described previously.[7] Baseline lumbar radiographs were performed at examinations 10 and 11 (1966 to 1970) as part of an osteoporosis evaluation. Participants were included in the present analysis only if they were free of CVD (defined by the occurrence of stroke, transient ischemic attack, congestive heart failure, IC, angina pectoris, coronary insufficiency, or myocardial infarction) at the x-ray examination.[8] Subjects were followed for 32 years for the development of IC, IS, CHD, and total CVD outcomes. We previously reported associations of AAC with CHD and total CVD.[2,9] The present work incorporates an additional 10 years of prospective follow-up of CHD and CVD end points. All FHS subjects provided written informed consent to participate in the study.

The scoring method to quantify lumbar x-ray calcification of the abdominal aorta has been previously described.[10] AAC was quantified in the first through fourth lumbar vertebrae. All x-rays were evaluated by an independent reader blinded to participant's clinical status. At each vertebra, the anterior and posterior aortic walls were assigned a score from 0 to 3 depending on the extent of calcification and the scores summed, resulting in a maximum total possible score of 24. A score of 0 denoted no calcification, 1 indicated small scattered deposits that filled $<1/3$ of the wall of the aorta, 2 indicated calcification in $>1/3$ but $<2/3$ of the aortic wall, and 3 indicated calcification in $>2/3$ of the aortic wall.[2,9] We previously reported excellent interrater reliability (0.93) and intrarater reliability (0.98) for this scoring method.[10]

Medical histories were elicited and physical examinations performed at biennial visits, as previously described.[11] Blood pressure was measured 2 times in the left arm, after the participant had been seated for $\geq$5 minutes. The average of 2 readings was used for analyses. All participants had blood samples drawn at the time of the examination, which included serum glucose, total cholesterol, and high-density lipoprotein cholesterol. Diabetes mellitus was defined as a random glucose level $\geq$11.1 mmol/L (200 mg/dl), a fasting serum glucose level $>$7.8 mmol/L ($>$140 mg/dl), or use of insulin or oral hypoglycemic agents.[12,13] Although the American Diabetes Association guidelines for diagnosing diabetes have changed significantly in recent years, these previous cutpoints were used to maintain consistency with previously published reports using the same cohort.[12,14] Left ventricular hypertrophy on electrocardiogram was diagnosed if a subject met voltage criteria that occurred concomitantly with lateral repolarization changes.[14] Body mass index was calculated as weight in kilograms divided by height in meters squared. Current cigarette smoking was identified in participants who reported smoking cigarettes regularly within 1 year of the examination.

Biennial FHS examinations and or health history updates were used to identify incident cases of IC, IS, CHD, and other forms of CVD. For IC, a physician-administered standardized questionnaire was used to elicit subjective symptoms of calf discomfort with exertion that occurred sooner with uphill or fast-paced walking and was alleviated with rest.[8,15] All suspected claudication events were verified independently by a second physician examiner.

Continuous surveillance for cerebrovascular events included daily hospital monitoring, tracking of all medical encounters, and in-person examination by an FHS stroke neurologist of all participants with possible stroke symptoms. All events were adjudicated by a panel of $\geq$2 neurologists and verification of stroke was available in 90% of stroke cases using imaging. Stroke occurrence and characteristics, including subtypes, were determined at the end of the acute phase of stroke using uniform criteria and a standardized protocol.[16]

Component events defining CHD (myocardial infarction, coronary insufficiency, angina, and coronary disease death) have been previously reported.[17] Definite congestive heart failure was defined by presence of 2 major criteria or 1 major criterion and 2 minor criteria as described in previous Framingham reports.[8,18] All CVD outcomes were adjudicated by a committee of 3 physicians using information from FHS physician examinations and all available records from the participant's personal physician and hospital medical records. Noninvasive testing (such as ankle-brachial index) was not used to confirm the diagnosis of occlusive peripheral arterial disease because it was not available throughout most of the follow-up period for this study.

Analyses were performed by tertiles of AAC scores as defined previously.[2,9] Gender-specific and gender-pooled Cox proportional hazard regression analyses were used to calculate age- and multivariable-adjusted hazard ratios (HRs) for development of CVD outcomes over 32 years of follow-up. Subjects in the lowest tertile (with an AAC score of 0) served as the reference group. Covariates in multivariable analysis included age, gender, diabetes, systolic blood pressure, left ventricular hypertrophy on electrocardiogram, body mass index, total cholesterol, high-density lipoprotein cholesterol, current cigarette smoking, and hypertension treatment. In a secondary analysis, a multivariable-adjusted model incorporated the same covariates in addition to intercurrent CVD.

To examine whether AAC contributed to prediction of cardiovascular outcomes over and above traditional CHD risk factors in patients at intermediate risk, a complementary analysis was performed examining each outcome separately (IC, IS, CHD, and CVD) while stratifying by the Framingham CHD risk score, using guidelines published in 1998.[19] Participants were stratified into clinically relevant 10-year CHD risk categories as previously described (low risk $<$6%, intermediate risk 6% to 20%, high risk $>$20%).[20,21] Cumulative incidence was calculated for each end point because $<$4% of subjects have been lost to follow-up for these events.

Discrimination, a model's ability to correctly distinguish events and nonevents on the basis of the baseline risk factor profile, was calculated using the c-statistic.[22] Reported c-statistic levels range from 0.50 (no discrimination) up to a maximum of 1.0 (perfect discrimination). Comparisons were made between the risk factor-adjusted models with AAC and without AAC. Statistical methods included Cox proportional hazards regression and discrimination analyses that included c-statistic comparisons with confidence intervals (CIs) around the estimates using resampling methods.[23]

All statistical analyses were performed using SAS software (SAS Institute, Cary, North Carolina).

Table 1
Baseline characteristics of study sample (those with abdominal aortic x-rays, n = 2,149)

Characteristic	Men (n = 848)	Women (n = 1,301)
Age (yrs)	60 ± 8	60 ± 8
Cholesterol (mg/dl)	220 ± 40	241 ± 41
High-density lipoprotein cholesterol (mg/dl)	45 ± 13	58 ± 16
Systolic blood pressure (mm Hg)	138 ± 21	139 ± 24
Diastolic blood pressure (mm Hg)	82 ± 11	80 ± 11
Blood pressure treatment (%)	10%	15%
Current smoking (%)	37%	34%
Diabetes mellitus (%)	4%	4%
Body mass index (kg/m^2)	27 ± 4	26 ± 4
Left ventricular hypertrophy (%)	1.7%	1%
AAC index	3 ± 4	3 ± 4

Values are means ± SDs or percentages of patients.

Table 2
Abdominal aortic calcium tertiles in men and women

Tertile	AAC Score	Men	Women	Total
1	0	302 (14%)	580 (27%)	882 (41%)
2	1–3	247 (12%)	286 (13%)	533 (25%)
3	4–24	299 (14%)	435 (20%)	734 (34%)
Total		848 (40%)	1,301 (60%)	2,149 (100%)

Results

There were 2,149 participants in the study sample (61% women, mean age 59.9 years). Baseline characteristics of the sample are listed in Table 1. The AAC score was skewed with about 1/3 having 0 scores, so AAC was classified in tertiles as previously described (Table 2).[2,9]

Age-adjusted incidence of IC over the course of 32 years of follow-up in men and women is displayed in Figure 1. In age-adjusted models, there was a graded increase in risk for IC in men and women from the first through third tertiles of AAC (Table 3). In men and women combined, the age-adjusted HR for IC in the third compared with the first AAC tertile was 2.54 (95% CI 1.79 to 3.61). After multivariable adjustment, the HR attenuated to 1.68 (95% CI 1.12 to 2.50) but remained statistically significant. In gender-specific analyses, the HR for IC in the third AAC tertile was significant for men and women, although the magnitude of risk was lower and of marginal statistical significance after multivariable adjustment (Table 3). In a secondary analysis adjusting for the impact of intercurrent CVD other than IC over the 32-year follow-up in addition to the other covariates in the multivariable model, the HR for the combined cohort was similar.

In men and women combined, the risk of IS was significantly increased in the third AAC compared with the first AAC tertile in age-adjusted (HR 2.17, 95% CI 1.49 to 3.14) and multivariable-adjusted (HR 1.73, 95% CI 1.12 to 2.65) models (Table 3). In women, there was a clear, graded increase in risk from the first to third tertiles of AAC; the multivariable-adjusted HR for the third tertile was 2.29 (95% CI 1.30 to 4.02). In men, age-adjusted risks increased across tertiles of AAC, although the magnitudes of the HRs became small and not statistically significant in multivariable-adjusted models. Similar findings were observed in models adjusting for intercurrent CVD.

In the combined cohort (Table 3), the multivariable HR for CVD for the third compared with the first tertile of AAC was 1.64 (95% CI 1.37 to 1.97) with comparable results in gender-specific multivariable analyses. Similar grade, increasing, statistically significant associations of AAC with outcome were also noted for CHD in the combined cohort and in gender-specific analysis, although the multivariable-adjusted HR for men was of borderline statistical significance (Table 3).

Cumulative incidence by tertile of AAC for each end point stratified by FHS 10-year CHD risk score for men and women combined are displayed in Figure 1. About 1/2 (53%, n = 1143) of our subjects were in the intermediate risk category, and 30% of intermediate-risk subjects (16% overall) were in the highest tertile of AAC. There was a consistently significant increase in IC, IS, CHD, and CVD incidence across tertiles of AAC in the low- and intermediate-risk groups (all p values <0.05), whereas risks tended to be not significantly different across tertiles of AAC in the highest-risk group (Figure 1). Incidence in the highest tertile of AAC in intermediate-risk subjects was generally as high as or higher than that in the first tertile of AAC in high-risk subjects. In gender-specific analyses, these differences were noted in women but not in men. For example, for IC in the combined cohort (Figure 1), the event rate in the third tertile of AAC in the low-risk group (108.5 per 1,000) and intermediate-risk group (115.5 per 1,000) was higher than the event rate for the first tertile of AAC in the FHS high-risk group (94.2 per 1,000). Similar patterns of association were noted for IS, CHD, and CVD. In women alone, incidence for all end points in low- and intermediate-risk groups in the third tertile of AAC consistently exceeded the incidence for high-risk participants in the first tertile of AAC.

To address whether AAC added to risk prediction for IC, IS, CHD, and CVD beyond traditional cardiovascular risk factors, c-statistics at 10 years of follow-up were calculated first for the multivariable-adjusted models and then for the multivariable-adjusted model including AAC (Table 4).[22] In the combined cohort, AAC added modestly but significantly to the prediction of IS, CHD, and CVD. For example, for CHD the c-statistic improved from 0.690 to 0.705 (p = 0.004).

Discussion

We demonstrate that AAC is a strong predictor of increased risk of developing IC and IS, even after adjustment for traditional cardiovascular risk factors. The magnitude of risk conferred by AAC for IC and IS is comparable to or exceeds that for CHD or CVD. AAC determined by plain lumbar radiography appears to be a simple tool to improve risk stratification for patients at low or intermediate cardiovascular risk determined by previously published risk-factor algorithms.[19,24] Additionally, in discrimination analyses, AAC adds to the predictive model for IS, CVD, and CHD over and above traditional Framingham risk factors.

To our knowledge, only 1 other study has evaluated the relation between aortic calcium and symptomatic peripheral

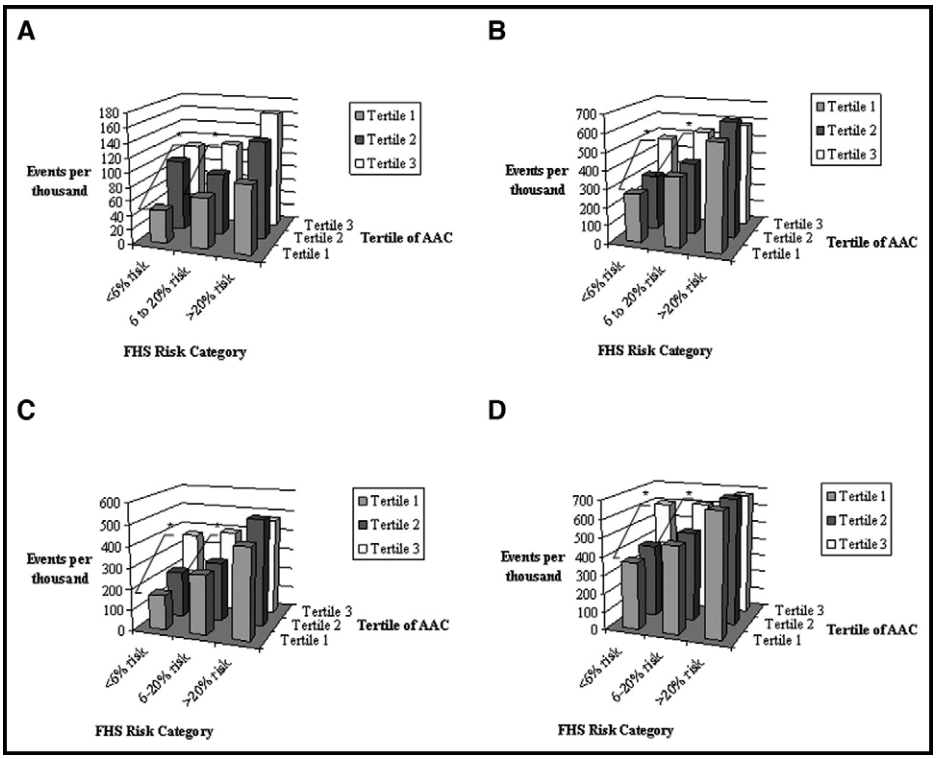


Figure 1. Age-adjusted event rate per 1,000 in men and women combined for IC *(A)*, IS *(B)*, CHD *(C)*, and total CVD *(D)* stratified by Framingham 10-year CHD risk score and tertile of AAC. *Significance at $p < 0.05$.

Table 3
Hazard ratios for intermittent claudication, ischemic stroke, coronary heart disease, and cardiovascular disease in men and women over 32 years of follow-up

	Model					
	Combined		Men		Women	
	Age	MV	Age	MV	Age	MV
IC						
Tertile 2 vs 1	1.64 (1.14–2.37)	1.15 (0.76–1.75)	1.29 (0.76–2.18)	1.18 (0.66–2.13)	2.11 (1.28–3.49)	1.17 (0.64–2.14)
Tertile 3 vs 1	2.54 (1.79–3.61)	1.68 (1.12–2.50)	2.14 (1.28–3.59)	1.68 (0.94–2.99)	2.97 (1.84–4.79)	1.75 (1.00–3.08)
No. at risk	2,149	1,605	848	646	1,301	959
No. events	199	159	94	78	105	81
IS						
Tertile 2 vs 1	1.86 (1.28–2.71)	1.44 (0.93–2.22)	1.38 (0.76–2.49)	1.06 (0.54–2.08)	2.26 (1.39–3.66)	1.76 (0.99–3.12)
Tertile 3 vs 1	2.17 (1.49–3.14)	1.73 (1.12–2.65)	1.69 (0.94–3.07)	1.11 (0.56–2.20)	2.51 (1.56–4.05)	2.29 (1.30–4.02)
No. at risk	2,149	1,605	848	646	1,301	959
No. events	201	150	78	58	123	92
CHD						
Tertile 2 vs 1	1.40 (1.15–1.71)	1.26 (1.01–1.58)	1.42 (1.09–1.86)	1.36 (1.00–1.84)	1.31 (0.97–1.76)	1.16 (0.82–1.65)
Tertile 3 vs 1	2.01 (1.66–2.44)	1.59 (1.26–2.00)	1.57 (1.18–2.09)	1.37 (0.99–1.90)	2.49 (1.91–3.24)	1.85 (1.34–2.56)
No. at risk	2,149	1,605	848	646	1,301	959
No. events	702	531	345	273	357	258
CVD						
Tertile 2 vs 1	1.42 (1.22–1.67)	1.27 (1.06–1.52)	1.34 (1.06–1.68)	1.34 (1.04–1.73)	1.48 (1.19–1.84)	1.23 (0.95–1.58)
Tertile 3 vs 1	2.00 (1.72–2.34)	1.64 (1.37–1.97)	1.75 (1.38–2.21)	1.55 (1.18–2.03)	2.20 (1.79–2.70)	1.80 (1.41–2.30)
No. at risk	2,149	1,605	848	646	1,301	959
No. events	1121	853	489	388	632	465

IC = intermittent claudication; IS = ischemic stroke; CHD = coronary heart disease; CVD = cardiovascular disease; MV = multivariable.

arterial disease, which found a significant relation with IC for women only.[25] Our finding that AAC is a strong predictor of incident stroke confirms and extends a previous report from the Rotterdam study.[26] In that report of older adults, aortic calcifications detected on lateral abdominal x-rays strongly predicted first-ever stroke and cerebral in-

Table 4
C-statistics for multivariable-adjusted models alone and multivariable-adjusted models with abdominal aortic calcium added at 10 years of follow-up

	Model					
	Combined		Men		Women	
	MV	MV + AAC	MV	MV + AAC	MV	MV + AAC
IC						
c-statistic	0.70	0.71	0.69	0.72	0.71	0.71
Δc-statistic	0.01		0.03		0.0006	
p value	0.10		0.37		0.48	
IS						
c-statistic	0.71	0.72	0.68	0.69	0.68	0.71
Δc-statistic	0.015		0.01		0.03	
p value	<0.01		0.08		<0.01	
CHD						
c-statistic	0.69	0.70	0.66	0.67	0.67	0.71
Δc-statistic	0.01		0.01		0.04	
p Value	<0.01		<0.05		<0.01	
CVD						
c-statistic	0.70	0.71	0.69	0.70	0.67	0.69
Δc-statistic	0.01		0.01		0.02	
p value	<0.01		<0.05		<0.05	

Δ = difference. Other abbreviation as in Table 3.

farction, even after adjustment for risk factors and other measurements of subclinical atherosclerosis. Although our sample was younger and free of baseline CVD, our multivariable-adjusted HRs were similar to those reported in the Rotterdam study. Thus, the atherosclerotic burden detected as AAC in the abdominal aorta appears to provide a potent index of cardiovascular risk in the lower extremity, cerebrovascular, and coronary arterial beds.[27]

Our study using plain abdominal x-rays of the aorta suggests that AAC imaging, even using a simple noninvasive test such as plain radiography or lateral imaging using dual x-ray absorptiometry, may have a useful role in risk prediction. In our sample, the magnitude of risk predicted by AAC for IC was stronger than that for CHD and for CVD in age- and multivariable-adjusted models. Similar magnitudes of risk were observed for IS. The HRs for CHD are similar but not identical to those previously reported from Framingham; the present HRs vary slightly from those previously reported because we include data from an additional 10 years of follow-up available for this study.[2,9] Overall, AAC adds to the prediction for IC, IS, CHD, and CVD over and above traditional CHD risk factors in low- and intermediate-risk patients, particularly women. In discrimination analyses, AAC adds significantly to prediction of IS, CHD, and CVD at 10 years of follow-up over and above traditional cardiovascular risk factors, suggesting that AAC may have utility as a screening tool. Thus, AAC appears to be a marker of global atherosclerosis and a strong predictor of symptomatic disease in multiple arterial beds. Our findings suggest that detection of AAC may be clinically useful for risk stratification in patients at intermediate risk, in whom decisions are often uncertain regarding implementation of risk factor modification.

Several limitations warrant consideration. First, some calcific plaques may result from conditions other than atherosclerosis, such as systemic diseases associated with vascular calcification, like end-stage renal disease. However, in a previous necropsy study of patients who also had lumbar x-rays to detect AAC, quantified using a protocol similar to ours, the vast majority of patients with AAC had grade 3 advanced atherosclerosis with calcification in the tunica intima.[28] Second, our data are drawn from a community-based sample beginning in the late 1960s that was largely untreated for CVD and risk factors by current standards. Although this provides an exceptional perspective on the natural history of atherosclerosis, the present population may carry less atherosclerotic burden due to more aggressive risk factor modification. Third, for detection of peripheral arterial disease we would have ideally used IC events or low ankle-brachial index as an end point; however, the ankle-brachial index was not available as a research tool in 1966 to 1970. Fourth, the Framingham cohort is predominantly white, and similar studies are warranted in subjects of other racial and ethnic backgrounds.

1. Pathobiological Determinants of Atherosclerosis in Youth (PDAY) Research Group. Natural history of aortic and coronary atherosclerotic lesions in youth. Findings from the PDAY Study. *Arterioscler Thromb* 1993;13:1291–1298.
2. Wilson PW, Kauppila LI, O'Donnell CJ, Kiel DP, Hannan M, Polak JM, Cupples LA. Abdominal aortic calcific deposits are an important predictor of vascular morbidity and mortality. *Circulation* 2001;103: 1529–1534.
3. Witteman JC, Kok FJ, van Saase JL, Valkenburg HA. Aortic calcification as a predictor of cardiovascular mortality. *Lancet* 1986;2:1120–1122.
4. Iribarren C, Sidney S, Sternfeld B, Browner WS. Calcification of the aortic arch: risk factors and association with coronary heart disease, stroke, and peripheral vascular disease. *JAMA* 2000;283:2810–2815.
5. Witteman JC, Kannel WB, Wolf PA, Grobbee DE, Hofman A, D'Agostino RB, Cobb JC. Aortic calcified plaques and cardiovascular disease (the Framingham Study). *Am J Cardiol* 1990;66:1060–1064.
6. Oei HH, Vliegenthart R, Hak AE, Iglesias dS, Hofman A, Oudkerk M, Witteman JC. The association between coronary calcification assessed by electron beam computed tomography and measures of extracoronary atherosclerosis: the Rotterdam Coronary Calcification Study. *J Am Coll Cardiol* 2002;39:1745–1751.

7. Dawber T, Kannel WB, Lyell L. An approach to longitudinal studies in a community: the Framingham Study. *Ann N Y Acad Sci* 1963;107: 539–556.
8. Cupples L, D'Agostino, Kiely D. The Framingham Heart Study, Section 35: An Epidemiological Investigation of Cardiovascular Events: Survival Following Initial Cardiovascular Events: 30 Year Follow-Up. Bethesda, MD: National Heart, Lung, and Blood Institute, 1988.
9. Walsh CR, Cupples LA, Levy D, Kiel DP, Hannan M, Wilson PW, O'Donnell CJ. Abdominal aortic calcific deposits are associated with increased risk for congestive heart failure: the Framingham Heart Study. *Am Heart J* 2002;144:733–739.
10. Kauppila LI, Polak JF, Cupples LA, Hannan MT, Kiel DP, Wilson PW. New indices to classify location, severity and progression of calcific lesions in the abdominal aorta: a 25-year follow-up study. *Atherosclerosis* 1997;132:245–250.
11. Feinleib M, Kannel WB, Garrison RJ, McNamara PM, Castelli WP. The Framingham Offspring Study. Design and preliminary data. *Prev Med* 1975;4:518–525.
12. National Diabetes Data Group. Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. *Diabetes* 1979; 28:1039–1057.
13. 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* 2003;289:2560–2572.
14. Kannel WB, Gordon T, Offutt D. Left ventricular hypertrophy by electrocardiogram. Prevalence, incidence, and mortality in the Framingham study. *Ann Intern Med* 1969;71:89–105.
15. Murabito JM, D'Agostino RB, Silbershatz H, Wilson WF. Intermittent claudication. A risk profile from the Framingham Heart Study. *Circulation* 1997;96:44–49.
16. Adams HP Jr, Bendixen BH, Kappelle LJ, Biller J, Love BB, Gordon DL, Marsh EE III. Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment. *Stroke* 1993;24:35–41.
17. Abbot R, McGee D. The probability of developing certain cardiovascular diseases in eight years at specified values of some characteristics. In: The Framingham Study: An Epidemiological Investigation of Cardiovascular Disease, Section 37. Bethesda, MD: National Heart, Lung, and Blood Institute; 1987.
18. McKee PA, Castelli WP, McNamara PM, Kannel WB. The natural history of congestive heart failure: the Framingham study. *N Engl J Med* 1971;285:1441–1446.
19. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation* 1998;97:1837–1847.
20. Can atherosclerosis imaging techniques improve the detection of patients at risk for ischemic heart disease? Proceedings of the 34th Bethesda Conference. Bethesda, Maryland, USA. October 7, 2002. *J Am Coll Cardiol* 2003;41:1856–1917.
21. Greenland P, Smith SC Jr, Grundy SM. Improving coronary heart disease risk assessment in asymptomatic people: role of traditional risk factors and noninvasive cardiovascular tests. *Circulation* 2001;104: 1863–1867.
22. 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.
23. Antolini L, Nam B-H, D'Agostino RB. Inference on correlated discrimination measures in survival analysis: a nonparametric approach. *Commun Stat Theory Methods* 2004;33:2117–2135.
24. Third Report of the National Cholesterol Education Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (ATP III) Final Report. NIH Publication Number 02-5215. Bethesda, MD: National Heart, Lung and Blood Institute, National Institutes of Health, 2002.
25. Danielsen R, Sigvaldason H, Thorgeirsson G, Sigfusson N. Predominance of aortic calcification as an atherosclerotic manifestation in women: the Reykjavik study. *J Clin Epidemiol* 1996;49:383–387.
26. Hollander M, Hak AE, Koudstaal PJ, Bots ML, Grobbee DE, Hofman A, Witteman JC, Breteler MM. Comparison between measures of atherosclerosis and risk of stroke: the Rotterdam Study. *Stroke* 2003; 34:2367–2372.
27. van der Meer, I, Bots ML, Hofman A, del Sol AI, van der Kuip DA, Witteman JC. Predictive value of noninvasive measures of atherosclerosis for incident myocardial infarction: the Rotterdam Study. *Circulation* 2004;109:1089–1094.
28. Hyman J, Epstein F. A study of the correlation between roentgenographic and post-mortem calcification of the aorta. *Am Heart J* 1954; 48:540–543.

Journal of the American College of Cardiology
© 2008 by the American College of Cardiology Foundation
Published by Elsevier Inc.

Vol. 51, No. 3, 2008
ISSN 0735-1097/08/$34.00
doi:10.1016/j.jacc.2007.09.038

Impact of Impaired Fasting Glucose on Cardiovascular Disease

The Framingham Heart Study

Yamini S. Levitzky, MD,* Michael J. Pencina, PHD,*‡ Ralph B. D'Agostino, PHD,*‡ James B. Meigs, MD, MPH,§ Joanne M. Murabito, MD, SCM,* ** Ramachandran S. Vasan, MD,*‖# Caroline S. Fox, MD, MPH*† ††

Framingham and Boston, Massachusetts; and Bethesda, Maryland

Objectives We sought to determine whether impaired fasting glucose (IFG) predicts cardiovascular disease (CVD) events.

Background It is unclear which glucose threshold should define prediabetes. We compared the 1997 and 2003 American Diabetes Association (ADA) definitions of IFG to predict CVD.

Methods Framingham offspring participants free of CVD, categorized by the 1997 ADA IFG definition (fasting plasma glucose 110 to 125 mg/dl; 6.1 to 6.9 mmol/l) or the 2003 definition (100 to 125 mg/dl; 5.6 to 6.9 mmol/l), were followed from 1983 to 2004. Pooled logistic regression was used to calculate multivariable-adjusted odds ratios (ORs) for incident coronary heart disease (CHD; 291 events) or CVD (423 events).

Results Four-year CHD event rates among women were 1.3% (100 to 109 mg/dl), 2.3% (110 to 125 mg/dl), and 2.9% (diabetes); whereas corresponding rates in men were 2.9%, 3.0%, and 8.7%. For the 2003 IFG definition, the OR for CHD among women was 1.7 (95% confidence interval [CI] 1.0 to 3.0, p = 0.048), whereas for the 1997 IFG definition, the OR for CHD in women was 2.2 (95% CI 1.1 to 4.4, p = 0.02), which was almost as high as for women with diabetes (OR 2.5, 95% CI 1.2 to 5.2, p = 0.01). For CVD, only the 1997 IFG definition yielded significantly greater odds of CVD in women (OR 2.1, 95% CI 1.2 to 3.6, p = 0.01). Men were not at increased odds of developing CVD or CHD by either definition.

Conclusions In women, both IFG definitions were associated with increased CHD risk, whereas neither IFG definition identified men at increased short-term risk for CHD or CVD. The finding that women with FPG 110 to 125 mg/dl had similar CHD risk compared with women with diabetes suggests that CHD risk in women may be elevated at a lower glucose level than for men. (J Am Coll Cardiol 2008;51:264–70) © 2008 by the American College of Cardiology Foundation

It has been recognized that prediabetic hyperglycemia confers an increased risk for cardiovascular disease (CVD) (1,2). In 1997, the American Diabetes Association (ADA) introduced the concept of impaired fasting glucose (IFG), a prediabetic state initially defined by fasting plasma glucose (FPG) of 110 to 125 mg/dl (6.1 to 6.9 mmol/l), in which those afflicted were significantly more likely to develop diabetes (3–5). The risk of developing CVD was not considered in establishing criteria for IFG.

Since the introduction of the concept of IFG, there has been considerable debate regarding where the lower limit should be set to achieve a reasonable balance between sensitivity and specificity for diabetes prediction. In 2003, the ADA lowered its threshold for diagnosis of IFG from 110 mg/dl (6.0 mmol/l) to 100 mg/dl (5.6 mmol/l) on the basis of evidence in selected samples that suggested diabetes prediction may be optimized at a lower threshold (6). The effect of this lowered cut point is that a much larger proportion of the population is now considered to have IFG. Using data from the Third National Health and Nutrition Examination Survey, Benjamin et al. (7) found that the prevalence of IFG among adults was estimated to increase from 8.3% to 30.2%.

From the *National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, Massachusetts; †National Heart, Lung, and Blood Institute, Bethesda, Maryland; ‡Statistics and Consulting Unit, Department of Mathematics, Boston University, Boston, Massachusetts; §Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts; ‖Cardiology Section and the #Department of Preventive Medicine and Epidemiology, Boston University School of Medicine, Boston, Massachusetts; **Section of General Internal Medicine, Department of Medicine, Boston Medical Center, Boston, Massachusetts; and the ††Department of Endocrinology, Metabolism, and Diabetes, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts. This work was supported by the National Heart, Lung, and Blood Institute's Framingham Heart Study (N01-HC-25195). Dr. Meigs is supported by an American Diabetes Association Career Development Award. Dr. Vasan is supported by a Midcareer Investigator Award 2K24 HL4334.

Manuscript received May 24, 2007; revised manuscript received September 12, 2007, accepted September 17, 2007.

Since the publication of the 2003 IFG guidelines, relatively few studies have examined the impact of the 2003 IFG definition on CVD risk, and none have found a relation between FPG 100 to 125 mg/dl (5.6 to 6.9 mmol/l) and increased CVD risk or mortality (8–12). However, these studies have been limited by examination of samples with limited generalizability (9–11), relatively small samples with few CVD events during follow-up period (8), and potential inclusion of participants who develop diabetes in the IFG category (9,12).

Thus, on the basis of available data, it is uncertain whether the 2003 ADA definition of IFG offers improved risk prediction regarding cardiovascular disease as compared with the 1997 IFG definition. Therefore, the primary aims of this analysis were to characterize the new 2003 ADA definition of IFG in the Framingham Offspring study by examining incident CVD events as compared with the 1997 IFG definition. We also assessed the risk of developing diabetes based on the 2 IFG definitions.

Methods

Study sample. Participants for this study were drawn from the Framingham Offspring cohort. The design and inclusion criteria of the Framingham Heart Study have been described elsewhere (13). The current investigation included offspring participants who attended examinations (referred to as index examinations) in 1983 to 1987 (cycle 3), 1987 to 1991 (cycle 4), 1991 to 1995 (cycle 5), and 1995 to 1998 (cycle 6). Participants could contribute information to more than one examination cycle provided they reached the next examination cycle free of an outcome event of interest. All participants with CHD or CVD at the index examinations were excluded from further analyses. Participants were followed in approximately 4-year intervals, and events were accrued through December 31, 2004. Overall, 4,138 unique individuals contributed a total of 13,273 person-exams for analyses of incident CHD, and 4,058 unique individuals contributed 12,918 person-exams for analyses of incident CVD. For analyses involving incident diabetes, a total of 3,634 unique individuals free of diabetes and CHD at baseline were followed until diabetes or examination cycle 7 (1998 to 2001) contributing a total of 11,325 person-exams. The Institutional Review Board at Boston Medical Center approved the study protocol, and all participants gave written informed consent.

Baseline measurements and definitions. All Framingham clinic visits include a physician interview, physical examination, and laboratory tests. Participants who had a fasting plasma glucose ≥126 mg/dl (>7.0 mmol/dl) or were on insulin or oral hypoglycemic agents were considered to have diabetes. The 1997 ADA guidelines defined IFG as a FPG concentration of 110 to 125 mg/dl (6.1 to 6.9 mmol/l) (14), whereas the 2003 ADA guidelines define IFG as 100 to 125 mg/dl (5.6 to 6.9 mmol/l) (6).

Outcome ascertainment. The primary outcomes of interest were CHD, CVD, and diabetes. Coronary heart disease included cases of myocardial infarction, stable and unstable angina pectoris, and CHD death (15). Cardiovascular disease was defined as any CHD event, stroke, transient ischemic attack (TIA), intermittent claudication, congestive heart failure, or CVD death. Diabetes was defined as described previously in the previous section. A panel of 3 physicians reviewed each CHD and CVD event and adjudicated the end point according to pre-established criteria (16).

Abbreviations and Acronyms

ADA = American Diabetes Association
BMI = body mass index
CHD = coronary heart disease
CI = confidence interval
CVD = cardiovascular disease
FPG = fasting plasma glucose
IFG = impaired fasting glucose
IGT = impaired glucose tolerance
OR = odds ratio
TIA = transient ischemic attack

Covariates. Covariates were assessed and updated at all index examinations. Covariates included age, systolic blood pressure, hypertension treatment, total cholesterol to high-density lipoprotein cholesterol ratio, cigarette smoking within the past year, and body mass index (BMI). For incident diabetes, covariates were age, cigarette smoking within the past year, and BMI. Hypertension was defined as systolic blood pressure ≥140 mm Hg, diastolic blood pressure ≥90 mm Hg, or current treatment with antihypertensive medications. Current smoking was defined as at least 1 cigarette per day within 1 year of the index examination. Weight, measured to the nearest pound, was obtained with the participant wearing a gown without slippers or shoes. The BMI was calculated by dividing weight (kilograms) by square of height (meters^2).

Statistical analysis. A significant gender interaction was observed when age–gender-adjusted models were fit with an IFG-by-gender interaction. Therefore, all subsequent analyses were gender-specific.

All individuals with CHD or CVD at each index examination were excluded. Three separate models were used to examine incident CHD and CVD: 1) To examine the impact of the 1997 IFG definition on CHD and CVD risk, we compared FPG 110 to 125 mg/dl (6.1 to 6.9 mmol/l) to a referent group of FPG <110 mg/dl (<6.1 mmol/l). 2) To examine the impact of the 2003 IFG definition on CHD and CVD risk, we compared FPG 100 to 125 mg/dl (5.6 to 6.9 mmol/l) to a referent group of FPG <100 mg/dl (<5.6 mmol/l). 3) To directly compare the categorization and performance of the 1997 and 2003 IFG definitions, a multicategory model was created comparing both FPG 100 to 109 mg/dl (5.6 to 6.0 mmol/l) and 110 to 125 mg/dl (6.1 to 6.9 mmol/l) to a referent group of FPG <100 mg/dl (<5.6 mmol/l). This model ensured that the same referent group would be used to compare individuals in the 100 to 109 mg/dl category with those in the 110 to 125 mg/dl category.

JACC Vol. 51, No. 3, 2008
January 22, 2008:264–70

Table 1 Baseline Characteristics of Cohort*

Characteristic	Women (n = 2,163)	Men (n = 1,895)
Age, yrs	48 (10)	49 (10)
Glucose, mmol/l	5.1 (1.2)	5.4 (1.4)
Current smoker, %	29	28
Body mass index, kg/m^2	25.6 (5.4)	27.3 (3.9)
Systolic blood pressure, mm Hg	122 (18)	127 (16)
Hypertension treatment, %	14	16
Hypertension, %	26	35
Total/high-density lipoprotein cholesterol, ratio	4.0 (1.4)	5.0 (1.6)

Values are mean (SD) or percent. *Data represent unique individuals based on first exam attended and is based on the sample free of cardiovascular disease (n = 4,058).

For analyses of diabetes prediction, individuals with diabetes at each index examination were excluded and 3 similar models (1 to 3 as described in the previous paragraph) were constructed. Age-adjusted incidence rates of CHD, CVD, and diabetes were calculated for each FPG group (17). Pooled logistic regression was used to calculate the odds of developing CHD, CVD, or diabetes over the follow-up intervals using SAS version 9.1 (SAS Institute, Cary, North Carolina) (18). Pooled logistic regression has been shown to provide estimates similar to those generated from time-dependent Cox regression analysis (19).

Models were initially age-adjusted and then adjusted for covariates for each end point. To compare the predictive capacity of the 2003 versus 1997 IFG definition, we examined the c-statistics of all multivariable models. The c-statistic is a measure of model discrimination or concordance between the predictions and outcomes (20). Crude Kaplan-Meier curves were constructed using time to CVD stratified by glycemic category (FPG <100 mg/dl; 100 to 109 mg/dl; 110 to 125 mg/dl; diabetes).

Results

The overall sample (n = 4,058) consisted of 53% women, and the mean age was 49 years (Table 1).

Impact of the IFG definition on incident CHD and CVD. The 4-year rates of developing CHD are presented in Table 2. There were 291 cases of incident CHD. Four-year CHD event rates among women were 1.3% (100 to 109 mg/dl), 2.3% (110 to 125 mg/dl), and 2.9% (diabetes), whereas corresponding rates in men were 2.9%, 3.0%, and 8.7%. The odds ratio for CHD in women for the 2003 definition was 1.7 (95% confidence interval [CI] 1.0 to 3.0)

Table 2 4-Year Age-Adjusted Event Rates of CHD, CVD, and Diabetes by Baseline Glycemic Status Category

	Multicategory Model (mg/dl)			
Outcome	≤99	100 to 109	110 to 125	Diabetes
Any CHD event				
Women				
Events, n	39	14	12	13
Person-exams, n	5,563	882	365	318
4-year event rate (%) with 95% CI	0.8 (0.5–1.1)	1.3 (0.7–2.2)	2.3 (1.3–4.3)	2.9 (1.6–5.3)
Men				
Events, n	106	38	19	50
Person-exams, n	3,960	1,235	516	434
4-year event rate (%) with 95% CI	2.9 (2.3–3.7)	2.9 (2.0–4.0)	3.0 (1.9–4.8)	8.7 (6.4–11.7)
Any CVD event				
Women				
Events, n	71	19	18	20
Person-exams, n	5,457	850	352	299
4-year event rate (%) with 95% CI	1.4 (1.1–1.9)	1.8 (1.1–2.9)	3.6 (2.2–5.9)	4.8 (3.0–7.6)
Men				
Events, n	138	60	31	66
Person-exams, n	3,876	1,198	495	391
4-year event rate (%) with 95% CI	3.9 (3.2–4.8)	4.6 (3.5–6.0)	5.0 (3.4–7.1)	12.5 (9.7–16.0)
Diabetes				
Women				
Events, n	17	31	87	—
Person-exams, n	5,049	780	317	—
4-year event rate (%) with 95% CI	0.3 (0.2–0.6)	4.0 (2.7–6.0)	27.8 (21.6–35.1)	—
Men				—
Events, n	23	52	92	—
Person-exams, n	3,608	1,129	442	—
4-year event rate (%) with 95% CI	0.6 (0.4–1.0)	4.5 (3.3–6.1)	20.0 (15.8–25.1)	—

CHD = coronary heart disease; CI = confidence interval; CVD = cardiovascular disease.

Table 3 Odds Ratios and 95% Confidence Intervals by Glucose Subgroup for 4-Year Incidence of CHD According to 1997 and 2003 IFG Definitions

	Multicategory Model (mg/dl)*						2003 IFG (mg/dl)*		1997 IFG† (mg/dl)	
	100 to 109	p Value	110 to 125	p Value	Diabetes	p Value	100 to 125	p Value	110 to 125	p Value
Women										
Age-adjusted	1.7 (0.9–3.1)	0.11	3.1 (1.6–6.1)	0.001	4.0 (2.1–7.6)	<0.001	2.1 (1.3–3.5)	0.005	2.7 (1.4–5.2)	0.002
MV-adjusted‡	1.4 (0.7–2.7)	0.30	2.5 (1.2–5.0)	0.01	2.5 (1.2–5.2)	0.01	1.7 (1.0–3.0)	0.048	2.2 (1.1–4.4)	0.02
Men										
Age-adjusted	1.0 (0.7–1.4)	0.92	1.0 (0.6–1.7)	0.91	3.2 (2.2–4.6)	<0.001	1.0 (0.7–1.4)	0.99	1.0 (0.6–1.7)	0.89
MV-adjusted‡	0.9 (0.6–1.4)	0.66	0.9 (0.5–1.5)	0.60	2.6 (1.7–3.8)	<0.001	0.9 (0.6–1.3)	0.55	0.9 (0.5–1.5)	0.67

*Referent group is <100 mg/dl (<5.6 mmol/l). †Referent group is <110 mg/dl (<6.1 mmol/l). ‡Covariates: age, systolic blood pressure, hypertension treatment, total cholesterol/high-density lipoprotein ratio, current smoking, and body mass index.

CHD = coronary heart disease; IFG = impaired fasting glucose; MV = multivariable.

and for the 1997 definition was 2.2 (95% CI 1.1 to 4.4) (Table 3). In the multicategory model, women with FPG 110 to 125 mg/dl (6.1 to 6.9 mg/dl) had a 2.5-fold increased odds ratio (OR) of CHD (95% CI 1.2 to 5.0, p = 0.01), whereas women with FPG 100 to 109 mg/dl (5.6 to 6.0 mmol/l) had a nonsignificantly increased OR of CHD (OR 1.4, 95% CI 0.7 to 2.7, p = 0.30), suggesting that much of the increase in CHD is driven primarily by those with FPG 110 to 125 mg/dl (6.1 to 6.9 mmol/l). Among women with FPG 110 to 125 mg/dl, the OR was similar to the OR for CHD in women with diabetes (OR 2.5, 95% CI 1.2 to 5.2, p = 0.01). The c-statistic was essentially unchanged between models based on the 1997 versus the 2003 IFG definition (0.798 vs. 0.800, respectively).

In men, there was no significant difference in the OR of developing CHD among those categorized by either the 1997 IFG definition or the 2003 IFG definition (Table 3). There was a significant gender interaction with incident CHD (p = 0.04), indicating that the OR of developing CHD in women was significantly greater than for men.

Similar trends were observed for CVD (Table 4) (423 incident events). The OR for CVD in women for the 2003 definition was 1.4 (95% CI 0.9 to 2.1) and for the 1997 definition was 2.1 (95% CI 1.2 to 3.6) (Table 4). In the multicategory model for women, no significantly increased OR was observed in the 100 to 109 mg/dl (5.6 to 6.0 mmol/l) group, whereas significantly increased OR were observed among women with FPG 110 to 125 mg/dl (6.1 to 6.9 mmol/l). The multivariable-adjusted c-statistics for the 1997 and 2003 definitions were similar (0.785 vs. 0.787, respectively).

In men, there was no significant difference in the OR of developing CVD in those with IFG as categorized by either the 1997 definition or 2003 definition; similar results were observed in the multicategory model (Table 4). The multivariable-adjusted c-statistics were identical for the 1997 and 2003 definitions (0.763). The formal sex-interaction was not significant (p = 0.19), although the trends observed were overall similar to the CHD results. The results are presented as Kaplan-Meier curves in Figure 1.

Impact of the IFG definition on incident diabetes. The 4-year age-adjusted rates of developing diabetes by glycemic category are presented in Table 2. Among women, the OR of developing diabetes was elevated in multivariable-adjusted models for both the 1997 and 2003 definitions (Table 5). Among men, the differences between the 1997 and 2003 definitions were less striking than in women (Table 5).

Secondary analyses. When glucose was modeled as a continuous variable, for CVD, significant results were observed for women (p = 0.007) as well as for men (p = 0.0001). However, when individuals with diabetes were excluded from the models, results were no longer significant (women p = 0.16; men p = 0.89). For CHD, similar findings were observed (data not shown). Additional adjustment for triglycerides in men did not result in any material changes to the results (data not shown).

Table 4 Odds Ratios and 95% Confidence Intervals by Glucose Subgroup for 4-Year Incidence of Cardiovascular Disease According to 1997 and 2003 IFG Definitions

	Multicategory Model, mg/dl*						2003 IFG, mmol/l*		1997 IFG, mmol/l†	
	100 to 109	p Value	110 to 125	p Value	Diabetes	p Value	100 to 125	p Value	110 to 125	p Value
Women										
Age-adjusted	1.3 (0.8–2.1)	0.38	2.6 (1.5–4.5)	<0.001	3.6 (2.1–6.0)	<0.001	1.7 (1.1–2.5)	0.01	2.5 (1.5–4.2)	<0.001
MV-adjusted‡	1.1 (0.6–1.8)	0.84	2.1 (1.2–3.7)	0.01	2.3 (1.3–4.1)	0.007	1.4 (0.9–2.1)	0.16	2.1 (1.2–3.6)	0.01
Men										
Age-adjusted	1.2 (0.9–1.6)	0.32	1.3 (0.8–1.9)	0.25	3.6 (2.6–5.0)	<0.001	1.2 (0.9–1.6)	0.19	1.2 (0.8–1.8)	0.35
MV-adjusted‡	1.1 (0.8–1.5)	0.51	1.0 (0.7–1.6)	0.85	2.8 (2.0–4.0)	<0.001	1.1 (0.8–1.4)	0.56	1.0 (0.7–1.5)	0.98

*Referent group is <5.6 mmol/l (<100 mg/dl). †Referent group is <6.1 mmol/l (<110 mg/dl). ‡Covariates: age, systolic blood pressure, hypertension treatment, total cholesterol/HDL ratio, current smoking, and BMI.

Abbreviations as in Table 3.

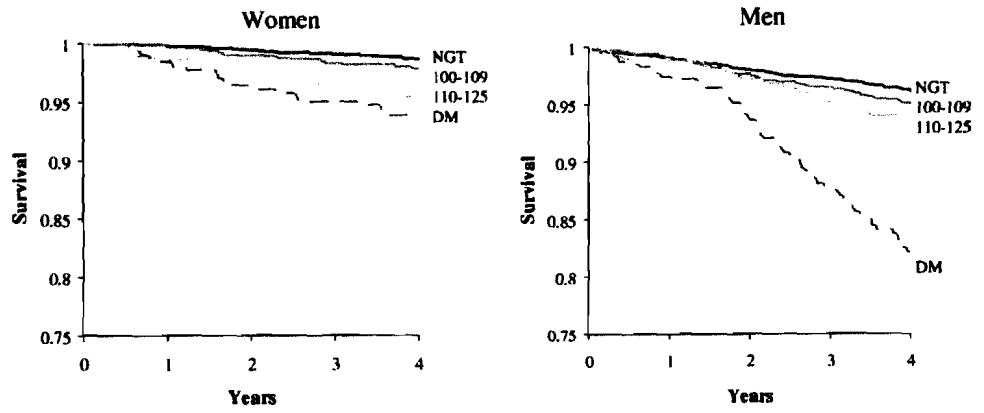


Figure 1 **Kaplain-Meier Plots of CVD-Free Survival by Glycemic Category**

CVD = cardiovascular disease; DM = diabetes mellitus; NGT = normal glucose tolerance; 100–109 = 100 to 109 mg/dl; 110–125 = 110 to 125 mg/dl.

Discussion

Principal findings. The 1997 and 2003 IFG definitions are predictive of CHD in women but not in men. The odds of developing CHD among women with IFG in the 110 to 125 mg/dl (6.1 to 6.9 mmol/l) range approach the risk conferred by having diabetes. C-statistics were essentially unchanged between the 1997 and 2003 IFG definitions, suggesting no improvement in overall risk prediction when one uses the new IFG definition. For CVD, only the 1997 IFG definition was associated with a statistically significant increased odds among women, whereas no increased odds of CVD was observed in men for either IFG definition. Ultimately, however, men have greater absolute rates of events as compared with women. With respect to diabetes, the 1997 IFG definition is associated with a greater risk of developing diabetes compared with the 2003 IFG definition.

A possible explanation for the difference between CVD and CHD is that the CHD end point contains primarily hard diagnoses that are known to be highly associated with diabetes and pre-diabetes, such as myocardial infarction, whereas CVD contains end points which are potentially more heterogeneous, such as intermittent claudication and TIA. Nonetheless, women in our sample with FPG 110 to 125 mg/dl (6.1 to 6.9 mmol/l) are at significantly increased risk of both CHD and CVD.

In the context of current literature. Conflicting data exist regarding the effect of nondiabetic fasting hyperglycemia on cardiovascular risk. Whereas several studies have found that the 1997 IFG definition is associated with significantly increased risk for CVD (2,9,11,21), at least 5 studies have shown no significantly increased risk for CVD with the 1997 IFG definition (5,8,10,22,23).

Relatively fewer studies have examined the 2003 IFG cut point and its predictive capacity for CVD, and none have demonstrated an increased risk (8–10,24). Kanaya et al. (9) examined data from the Heart and Estrogen/Progestin Replacement Study, which enrolled women with known coronary disease, grouped them by fasting glucose status, and followed them for CVD events, stroke, TIA, and CHF hospitalization for an average of 6.8 years. They found that women with FPG 100 to 125 mg/dl (5.6 to 6.9 mmol/l) were at no increased risk for any end point as compared with women with normal levels of fasting glucose. In contrast, we found that women categorized by the 2003 IFG definition do not have a statistically significant increased risk of CVD

Table 5 **Odds Ratios and 95% Confidence Intervals for 4-Year Incidence of Diabetes Examining the 1997 and 2003 IFG Definitions in Age- and Multivariable-Adjusted Models**

	Multicategory Model, mmol/l				2003 IFG Glucose Category, mmol/l		1997 IFG Glucose Category, mmol/l	
	5.6 to 6.0*	p Value	6.1 to 6.9*	p Value	5.6 to 6.9*	p Value	6.1 to 6.9†	p Value
Women								
Age-adjusted	12.4 (6.8–22.8)	<0.001	114.5 (65.4–200.5)	<0.001	33.9 (20.1–57.3)	<0.001	42.3 (28.5–62.9)	<0.001
MV-adjusted‡	9.1 (4.9–17.0)	<0.001	72.5 (40.5–129.8)	<0.001	22.3 (13.0–38.1)	<0.001	26.3 (17.4–39.8)	<0.001
Men								
Age-adjusted	7.2 (4.4–11.9)	<0.001	38.5 (23.8–62.1)	<0.001	14.6 (9.3–22.8)	<0.001	14.9 (10.7–20.8)	<0.001
MV-adjusted‡	6.4 (3.9–10.6)	<0.001	32.4 (20.0–52.6)	<0.001	12.7 (8.1–20.0)	<0.001	12.9 (9.3–18.1)	<0.001

*Referent group is <5.6 mmol/l (<100 mg/dl). †Referent group is <6.1 mmol/l (<110 mg/dl). ‡Covariates: age, body mass index, smoking.
Abbreviations as in Table 3.

(OR 1.4, 95% CI 0.9 to 2.1) but do have significantly increased odds of CHD (OR 1.7, 95% CI 1.0 to 3.0, p = 0.048). A potential explanation for the differences in our findings may be the result of differences in our study sample, which included only individuals free of CVD at baseline, which is especially important when comparing our findings to those of Kanaya et al. (9), who used a sample of women with pre-existing CVD, and followed their participants for new events.

In a recent publication from the Hoorn Study in which participants (n = 1,428) were categorized according to 1997 and 2003 criteria based on OGTT measured in 1989 and 1996 with 10-year follow-up for all-cause and CVD mortality, there was no significant increased risk for CVD unless participants developed diabetes (12). These data are distinctly incongruent with our current findings, possibly in part because of the lack of gender-specific analyses in the Hoorn study.

The Hoorn investigators used oral glucose tolerance testing to diagnose impaired fasting glucose as compared with our use of FPG, which may also be pertinent to understanding why our findings differed. However, researchers using data from the Atherosclerosis Risk in Communities Study recently have confirmed that there is poor congruence between IFG (defined as 100 to 125 mg/dl; 5.6 to 6.9 mmol/l) and impaired glucose tolerance (IGT). They also demonstrate that neither IFG nor IGT are associated with an increased risk of all-cause mortality or incident CHD after a median follow-up of 6.3 years in fully-adjusted models (25). Therefore, we believe that our findings using FPG to diagnose IFG is an acceptable and clinically applicable method by which to conduct these analyses. Finally, a recent study from a community-based medical center examined CVD risk factor prevalence and prevalent CVD events among individuals with the 1997 as compared with the 2003 IFG definition (26) and found that the 2003 definition was not associated with an elevated level of CVD risk factors or CVD as compared to the 1997 definition.

Gender differences. For any given glycemic category, women had greater relative odds of CHD and CVD as compared with men, although men had greater absolute event rates for cardiovascular disease. In fact, the cardiovascular disease event rates and odds ratios for women in the 110 to 125 mg/dl group were similar to those for women with diabetes in our sample. These findings build upon those from a recent meta-analysis that included more than 33,000 women and 172,000 men in examining nondiabetic hyperglycemia as a risk factor for CVD; results demonstrate that the risk of CVD events was markedly greater in cohorts that included women (27). However, gender-specific data for women in this meta-analysis were not presented. Taken together with our findings, CVD and diabetes risk in women may occur at lower glucose thresholds as compared with men, which raises several potentially interesting questions. Whether gender differences are due to intrinsic biologic differences or differences in risk factor management is uncertain. These differences also raise the question of whether gender-specific cut points for impaired fasting glucose should exist.

Implications. In the absence of a clear glucose threshold that is predictive of CVD, the debate continues regarding what should define IFG to maximize sensitivity and specificity for predicting cardiovascular events. In examining the effect of IFG categorization with respect to cardiovascular disease, we uncovered gender differences, raising the question of whether a lower glycemic threshold should be used to diagnose IFG or diabetes in women. It is important to remember, however, that men have a greater absolute rate of events as compared with women. Further, IFG is not a CHD risk-equivalent. In addition, whether the effect of identifying individuals with this diagnosis in clinical practice encourages lifestyle modification including weight loss and increased physical activity is uncertain. Fasting blood sugar is often associated with other adverse CVD risk factors and may serve to identify patients with hypertension and dyslipidemia. Finally, it is uncertain whether identifying an individual with IFG results in aggressive CVD risk factor modification; randomized clinical trials among individuals with IFG would be necessary to assess this.

Strengths and study limitations. The strength of our analysis lies in our population-based cohort and long-term follow-up. We assessed the glycemic status of our participants every 4 years and were able to remove those who developed incident diabetes from the IFG category, which is particularly important when trying to understand the risk of cardiovascular disease associated with IFG independently of developing diabetes.

Our study has several limitations which must be acknowledged. Oral glucose tolerance testing was not available at each index examination cycle, thereby precluding comment on how the 1997 and 2003 IFG definitions compare with IGT for prediction of CVD. Further, this unavailability may have resulted in cases of undiagnosed diabetes in our exposure group. However, the primary point of our paper is to analyze the current ADA recommendations for IFG, and current guidelines do not recommend the routine use of oral glucose tolerance testing (28).

Next, although we found no increased risk in men for CHD and CVD, we examined short-term risk, and application to long-term risk, which allows for transition to diabetes over many years of follow-up, is uncertain. However, the consideration of short-term risk is congruent with the ADA's recommendation that all individuals over 45 years of age be screened for diabetes every 3 years and even more frequently if additional risk factors are present (29). Third, the Framingham Heart Study at its inception in 1948 included only white participants. Therefore, the generalizability to other ethnic groups is uncertain. However, the Framingham risk score has been validated in other ethic groups and has been found to be applicable in other populations (30,31). Although we used data from our study that spans several decades, we do not believe that temporal

trends would have an important effect on our results, because we have previously shown that the relative risk between cardiovascular disease and diabetes has not changed over time (32). Finally, we only evaluated cardiovascular complications of IFG and diabetes and were not able to assess retinopathy, neuropathy, and nephropathy, which may yield different findings.

Conclusions

Our data suggest that for prediction of CHD and CVD events, neither IFG definition identifies a group of men at increased short-term risk. In women, there is a significantly increased risk of CHD in the 2003 IFG group, but this risk is driven primarily by the high event rate in participants with FPG 110 to 125 mg/dl (6.1 to 6.9 mmol/l). In comparing women and men, for any given prediabetic category, women have a greater relative risk of CHD than men. The 2003 IFG definition does not offer substantive advantages over the 1997 definition for prediction of CVD or diabetes.

Reprint requests and correspondence: Dr. Caroline S. Fox, Framingham Heart Study, 73 Mount Wayte Avenue, Suite No. 2, Framingham, Massachusetts 01702-5827. E-mail: foxca@nhlbi.nih.gov.

REFERENCES

1. Wingard DL, Barrett-Connor EL, Scheidt-Nave C, McPhillips JB. Prevalence of cardiovascular and renal complications in older adults with normal or impaired glucose tolerance or NIDDM. A population-based study. Diabetes Care 1993;16:1022–5.
2. Coutinho M, Gerstein HC, Wang Y, Yusuf S. The relationship between glucose and incident cardiovascular events. A metaregression analysis of published data from 20 studies of 95,783 individuals followed for 12.4 years. Diabetes Care 1999;22:233–40.
3. Dinneen SF, Maldonado D III, Leibson CL, et al. Effects of changing diagnostic criteria on the risk of developing diabetes. Diabetes Care 1998;21:1408–13.
4. Chen KT, Chen CJ, Gregg EW, Imperatore G, Narayan KM. Impaired fasting glucose and risk of diabetes in Taiwan: follow-up over 3 years. Diabetes Res Clin Pract 2003;60:177–82.
5. de Vegt F, Dekker JM, Jager A, et al. Relation of impaired fasting and postload glucose with incident type 2 diabetes in a Dutch population: the Hoorn Study. JAMA 2001;285:2109–13.
6. Genuth S, Alberti KG, Bennett P, et al. Follow-up report on the diagnosis of diabetes mellitus. Diabetes Care 2003;26:3160–7.
7. Benjamin SM, Cadwell BL, Geiss LS, Engelgau MM, Vinicor F. A change in definition results in an increased number of adults with prediabetes in the United States. Arch Intern Med 2004;164:2386.
8. Blake DR, Meigs JB, Muller DC, Najjar SS, Andres R, Nathan DM. Impaired glucose tolerance, but not impaired fasting glucose, is associated with increased levels of coronary heart disease risk factors: results from the Baltimore Longitudinal Study on Aging. Diabetes 2004;53:2095–100.
9. Kanaya AM, Herrington D, Vittinghoff E, et al. Impaired fasting glucose and cardiovascular outcomes in postmenopausal women with coronary artery disease. Ann Intern Med 2005;142:813–20.
10. Tai ES, Goh SY, Lee JJ, et al. Lowering the criterion for impaired fasting glucose: impact on disease prevalence and associated risk of diabetes and ischemic heart disease. Diabetes Care 2004;27:1728–34.
11. Wen CP, Cheng TY, Tsai SP, Hsu HL, Wang SL. Increased mortality risks of pre-diabetes (impaired fasting glucose) in Taiwan. Diabetes Care 2005;28:2756–61.
12. Rijkelijkhuizen JM, Nijpels G, Heine RJ, Bouter LM, Stehouwer CD, Dekker JM. High risk of cardiovascular mortality in individuals with impaired fasting glucose is explained by conversion to diabetes: the Hoorn study. Diabetes Care 2007;30:332–6.
13. Dawber TR, Kannel WB, Lyell LP. An approach to longitudinal studies in a community: the Framingham Study. Ann N Y Acad Sci 1963;107:539–56.
14. Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Diabetes Care 1997;20:1183–97.
15. Cupples L, D'Agostino, Kiely D. The Framingham Heart Study, Section 35: An Epidemiological Investigation of Cardiovascular Events: Survival Following Initial Cardiovascular Events: 30 Year Follow-Up. Bethesda, MD: National Heart, Lung, and Blood Institute, 1988.
16. Stokes J III, Kannel WB, Wolf PA, Cupples LA, D'Agostino RB. The relative importance of selected risk factors for various manifestations of cardiovascular disease among men and women from 35 to 64 years old: 30 years of follow-up in the Framingham Study. Circulation 1987;75:V65–73.
17. Chang IM, Gelman R, Pagano M. Corrected group prognostic curves and summary statistics. J Chronic Dis 1982;35:669–74.
18. SAS. User's Guide, version 9.1. Cary, NC: SAS Institute, 2006.
19. 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;9:1501–15.
20. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143:29–36.
21. Barzilay JI, Spiekerman CF, Wahl PW, et al. Cardiovascular disease in older adults with glucose disorders: comparison of American Diabetes Association criteria for diabetes mellitus with WHO criteria. Lancet 1999;354:622–5.
22. DECODE Study Group, the European Diabetes Epidemiology Group. Glucose tolerance and cardiovascular mortality: comparison of fasting and 2-hour diagnostic criteria. Arch Intern Med 2001;161:397–405.
23. Tominaga M, Eguchi H, Manaka H, Igarashi K, Kato T, Sekikawa A. Impaired glucose tolerance is a risk factor for cardiovascular disease, but not impaired fasting glucose. The Funagata Diabetes Study. Diabetes Care 1999;22:920–4.
24. Sorkin JD, Muller DC, Fleg JL, Andres R. The relation of fasting and 2-h postchallenge plasma glucose concentrations to mortality: data from the Baltimore Longitudinal Study of Aging with a critical review of the literature. Diabetes Care 2005;28:2626–32.
25. Pankow JS, Kwan DK, Duncan BB, et al. Cardiometabolic risk in impaired fasting glucose and impaired glucose tolerance: the Atherosclerosis Risk in Communities Study. Diabetes Care 2007;30:325–31.
26. Kim SH, Chunawala L, Linde R, Reaven GM. Comparison of the 1997 and 2003 American Diabetes Association classification of impaired fasting glucose: impact on prevalence of impaired fasting glucose, coronary heart disease risk factors, and coronary heart disease in a community-based medical practice. J Am Coll Cardiol 2006;48:293–7.
27. Levitan EB, Song Y, Ford ES, Liu S. Is nondiabetic hyperglycemia a risk factor for cardiovascular disease? A meta-analysis of prospective studies. Arch Intern Med 2004;164:2147–55.
28. Report of the expert committee on the diagnosis and classification of diabetes mellitus. Diabetes Care 2003;26 Suppl 1:S5–20.
29. Standards of medical care in diabetes—2006. Diabetes Care 2006;29 Suppl 1:S4–42.
30. D'Agostino RB Sr., Grundy S, Sullivan LM, Wilson P. Validation of the Framingham coronary heart disease prediction scores: results of a multiple ethnic groups investigation. JAMA 2001;286:180–7.
31. Liu J, Hong Y, D'Agostino RB, et al. Predictive value for the Chinese population of the Framingham CHD risk assessment tool compared with the Chinese Multi-Provincial Cohort Study. JAMA 2004;291:2591–9.
32. Fox CS, Coady S, Sorlie PD, et al. Trends in cardiovascular complications of diabetes. JAMA 2004;292:2495–9.

BMC Public Health

Research article

Disclosing intimate partner violence to health care clinicians - What a difference the setting makes: A qualitative study

Jane Liebschutz*[1,2], Tracy Battaglia[1], Erin Finley[3,4] and Tali Averbuch[1]

Address: [1]Section of General Internal Medicine, Department of Medicine, Boston Medical Center and Boston University School of Medicine, 801 Massachusetts Ave, Boston, MA 02118, USA , [2]Department of Social and Behavioral Sciences, Boston University School of Public Health, 801 Massachusetts Ave, Boston, MA 02118, USA , [3]Department of Anthropology, Emory University, 115 Dickey Dr, Atlanta, GA 30322, USA and [4]The Emory Center for Myth and Ritual in American Life, Emory University, 115 Dickey Dr, Atlanta, GA 30322, USA

Email: Jane Liebschutz* - jane.liebschutz@bmc.org; Tracy Battaglia - tracy.battaglia@bmc.org; Erin Finley - epfinley@gmail.com; Tali Averbuch - tali.averbuch@gmail.com

* Corresponding author

Published: 4 July 2008

BMC Public Health 2008, **8**:229 doi:10.1186/1471-2458-8-229

Received: 14 March 2008
Accepted: 4 July 2008

This article is available from: http://www.biomedcentral.com/1471-2458/8/229

Abstract

Background: Despite endorsement by national organizations, the impact of screening for intimate partner violence (IPV) is understudied, particularly as it occurs in different clinical settings. We analyzed interviews of IPV survivors to understand the risks and benefits of disclosing IPV to clinicians across specialties.

Methods: Participants were English-speaking female IPV survivors recruited through IPV programs in Massachusetts. In-depth interviews describing medical encounters related to abuse were analyzed for common themes using Grounded Theory qualitative research methods. Encounters with health care clinicians were categorized by outcome (IPV disclosure by patient, discovery evidenced by discussion of IPV by clinician without patient disclosure, or non-disclosure), attribute (beneficial, unhelpful, harmful), and specialty (emergency department (ED), primary care (PC), obstetrics/gynecology (OB/GYN)).

Results: Of 27 participants aged 18–56, 5 were white, 10 Latina, and 12 black. Of 59 relevant health care encounters, 23 were in ED, 17 in OB/GYN, and 19 in PC. Seven of 9 ED disclosures were characterized as unhelpful; the majority of disclosures in PC and OB/GYN were characterized as beneficial. There were no harmful disclosures in any setting. Unhelpful disclosures resulted in emotional distress and alienation from health care. Regardless of whether disclosure occurred, beneficial encounters were characterized by familiarity with the clinician, acknowledgement of the abuse, respect and relevant referrals.

Conclusion: While no harms resulted from IPV disclosure, survivor satisfaction with disclosure is shaped by the setting of the encounter. Clinicians should aim to build a therapeutic relationship with IPV survivors that empowers and educates patients and does not demand disclosure.

Background

The extensive physical and mental health burden of intimate partner violence (IPV) exposure has been documented in various settings [1-6]. In response, the medical community has prioritized IPV identification. In fact, the Joint Commission on Accreditation of Healthcare Organ-

izations hospital standards call for IPV survivor identification and then referral to community services. In a complementary set of guidelines, the Family Violence Prevention Fund suggests clinicians inquire about IPV at every encounter for episodic care, such as Emergency Department visits, with higher case-finding rates as the measure of high quality [7].

Evidence to support IPV screening interventions includes surveys of patients who report expectations that a clinician inquire about IPV and increased satisfaction with the visit after being asked regardless of disclosure [8,9]. In a meta-analysis of the qualitative literature, Feder and colleagues found women IPV survivors value support and education whether or not they are ready to talk about the abuse, and consider most helpful relationships with physicians characterized by respectful support [10].

The most recent guidelines of the United States Preventive Services Task Force found insufficient evidence for screening for family violence due to lack of studies showing that a primary care based screening intervention helps reduce harmful outcomes [11]. In addition, the potential for negative outcomes of screening has not been examined [8,11-14]. Rather, the success of screening interventions tends to be measured in number of disclosures rather than in improvement of the survivor's overall condition [15-17]. Most studies of screening either lack a measure for the potential harm of disclosure or minimize such potential in written reports [15,17-19]. Although studies have shown higher disclosure in certain specialties [20], the difference in outcomes by specialty have not been well described. Rhodes and colleagues reported that many inquiries about IPV in an emergency room setting by physicians were perfunctory and did not lead to documentation or referral for other help [8,21].

Previously we reported results from a qualitative study of IPV survivors in which we examined those qualities of the patient-provider relationship that facilitate a safe and productive disclosure [22]. In that study, participants identified important provider characteristics, including: the ability to communicate a sense of personal concern; open communication; willingness to negotiate issues of control; confidentiality of medical information; shared decision-making; competency in medical care; careful listening; and taking ample time to address participant concerns [22]. Because of the reports of challenges to communication about IPV in emergency room settings and the emphasis on trust and communication in the patient-provider relationship, we hypothesized that the setting of disclosure of IPV might be important to the patient experience. Furthermore, such differences might inform clinical practice in varied medical settings.

In this paper we present the results of a re-analysis of participants' descriptions of patient-provider encounters to examine potential harms and benefits of IPV disclosure. We explored whether the specialty of care was related to the outcomes of disclosure, and identified a series of factors affecting these outcomes across primary care, obstetrics/gynecology and emergency department specialties.

Methods

Study Design

Ethnographic interviewing elicited IPV survivors' experiences interacting with both physician and non-physician health care providers. Grounded theory, a method of qualitative analysis [23], was used to elucidate views on patient-provider encounters revealed in the narrative data.

Study Participants

Twenty-seven IPV survivors were recruited from community-based domestic violence counseling or sheltering programs in eastern Massachusetts. They were recruited either through referral by local shelter staff or through a flier sent to all domestic violence programs in eastern Massachusetts. Eligible participants were female, ages 18 to 64, English-speaking, with a history of an abusive intimate partner relationship within the past 3 years. Each participant provided written informed consent and was compensated $25.

Interview Technique

After approval by the Boston University Medical Center Institutional Review Board, data were collected from October 1996 through November 2000. Open-ended, in-depth interviews, conducted by 1 of 2 authors (JL, TB), both primary care physicians, were audio-taped and lasted 1–2 hours.

Using an interview guide, the interviewer asked participants to describe encounters with health care clinicians both related and unrelated to the abusive relationship after the onset of the abuse. While most participants related to the onset of the adult intimate partner violence, others spontaneously mentioned experiences with health-care providers during adolescence or relating to childhood abuse. The participants were asked to provide information on perceived barriers to care and the abusive relationship over the past three years. Interviews were iterative; participants enrolled later in the data collection interval were questioned about themes revealed in previous interviews.

Analysis

Each audio-taped interview was transcribed verbatim by a professional transcriber, reviewed for accuracy and de-identified. Authors independently reviewed transcripts to identify common themes which were developed into a preliminary coding scheme with the first 10 interviews.

An advisory group of domestic violence advocates and survivors helped revise this scheme and suggest new concepts. The authors then independently coded the interviews using this revised coding scheme. Coding was compared and differences of opinion resolved through examination of the text.

Using NUD*ST qualitative research software (QSR International, Pty., Ltd., Melbourne, Australia) for data organization and coding, separate narratives representing a single patient-clinician relationship were identified and labeled as *encounters*. Encounters, which could be composed of a single interaction or continued contact over a period of years, were first categorized into "related to abuse" or "unrelated to abuse". As we did reiterative coding and analysis to understand the specific effect of disclosing (or not-disclosing) IPV, these unrelated encounters did not offer relevant material to allow categorization into a specific outcome (disclosure, discovery/discussion, nondisclosure) and were thus dropped from the analyses. Each medical encounter related to abuse was then coded according to three characteristics: outcome, specialty and attribute.

The first of these, outcome, described three mutually exclusive types of encounters: disclosure, discovery, and non-disclosure. A disclosure occurred when a participant reported telling her clinician about IPV. When a participant perceived her clinician knew of the abuse when she had not made an explicit disclosure, the outcome was labeled discovery. To be labeled discovery, the participants made explicit reference that the provider discussed some aspect of IPV, such as counseling or referral, even without explicit disclosure of IPV. All other encounters that did not fall into disclosure or discovery were labeled non-disclosure. To qualify for non-disclosure, one of two circumstances had to apply. First, the provider asked but the participant purposely did not disclose. Second, the participant was in an actively abusive relationship but did not spontaneously disclose, such as during treatment for injury, or during medical or pregnancy related care.

Each encounter was also coded for its *specialty*: Emergency Department (ED), Obstetrical or Gynecological Care (OB/GYN), Primary Care (PC) or other. PC included pediatricians and family physicians identified as the primary care provider but who may have also provided obstetrical care. Encounters occurring in other specialties (e.g. mental health, surgery) were excluded from this analysis because there were too few of any single type.

The final category, attribute, described the participant's level of satisfaction with the encounter as a result of whether she perceived the interaction as beneficial, harmful or unhelpful. For example, if an unpleasant interaction ended in the participant accepting help or receiving information that she found useful, we labeled it beneficial. Harmful interactions were ones resulting in injury to self, child or direct worsening of abuse. We classified negative reports not resulting in actual harm as unhelpful. When we were unable to categorize attribute due to a lack of information or contradictory descriptions, we excluded that encounter from analysis. Finally, we conducted a comparative analysis to explore the characteristics of encounters across outcomes, specialties and attributes.

Results

We interviewed 27 women; 12 were black, 10 Latina, and 5 White. Fourteen were recruited by domestic violence staff, and thirteen contacted the authors in response to the informational flier. Sixteen were living in a residential program at the time of the interview. Participant ages ranged from 18–56 years; median age was 31 years. Twenty-three participants had at least one child.

A total of 185 health care encounters were described. The number of encounters per participant ranged from 3–12; median number of encounters was 7. Although it was frequently difficult to determine the professional designation (physician, nurse, therapist) of an individual provider, specialty was clear in 175 encounters. The thirty-one mental health encounters were excluded because most were visits specifically related to the IPV. Twenty-two were nurses from different treatment settings (inpatient, public health, etc.). Of the twenty-nine other encounters, there were two few (<5) of any single type and could not be easily combined into categories- such as radiology technicians, surgeons, ambulance drivers, physical therapist, child protective service worker, medical subspecialist, etc. Thirty-one were excluded because they were unrelated to abuse, and did not contribute to the analysis presented in this paper, the impact of IPV disclosure. Another three were unable to be classified by attribute, leaving a sample pool of 59 encounters (23 primary care, 17 ED, 19 OB/GYN) representing 25 participants (Figure 1).

Thirty-five (59%) of these encounters involved IPV disclosure to the clinician, 7 (12%) in discovery, and 17 (29%) in non-disclosure. Of the disclosures, 25 (71%) were beneficial. Among discoveries, 4 were beneficial (57%), while among non-disclosures, 6 (35%) were beneficial. Setting of care was associated with reported satisfaction from disclosure. In the ED, 2 (22%) disclosures were beneficial. Of OB/GYN disclosures, 9 (75%) were beneficial. In primary care, all 14 disclosures were beneficial (Figure 2). There were no harmful disclosures in any specialty, and the remaining disclosures were unhelpful. We discuss these findings further in the paragraphs below.

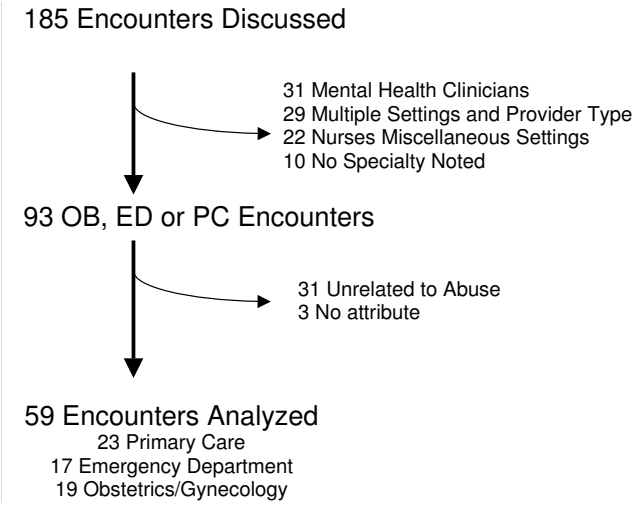


Figure 1
Encounter Classification Flow Chart.

Consequences of Unhelpful Disclosures: Fear and Avoidance of Healthcare

The most serious negative consequences of disclosure occurred in two participants who reported feeling endangered because of the disclosure, both after treatment for acute injuries in the ED. However, neither experienced any actual increase in violence. In two OB/GYN visits, participants found disclosure experiences so problematic they ultimately left their providers. The remaining 5 unhelpful disclosure experiences resulted in dissatisfaction without cessation of the clinical relationship.

Several participants were concerned by practitioners' tendency to encourage extreme "solutions" to the violence, like telling women to file a police report immediately. While in the ED, one participant reported being told, "Just tell me the name and where he's at and we'll send the police at him right now." She recalled thinking: "But what makes them think he won't come back and kill me?" The participant did not contact the police, and returned home in fear.

A number of participants indicated the cumulative effect of unhelpful disclosure experiences was avoidance of health care encounters. One participant noted:

> "*I used to go without medical treatment... I'd wait until it wasn't a choice anymore. And I'd wind up having to go to the emergency room.*"

Another participant revealed:

> "*Somebody would find out something was happening in my house, like a social worker, a doctor, a nurse or whatever, I would stop going there and go somewhere else.*"

The lack of an emotional connection with the clinician was a prominent feature in participants' discussion of unhelpful disclosures. Describing an ED visit, one participant commented:

> "*He checked me, he didn't ask any questions, nothing, and they took x-rays and pulled out of there... Maybe I was hoping... that they would talk to me? I mean, they checked me out... but I didn't feel like... emotionally? Like maybe talk, some kind of comfort?*"

The lack of effective communication on safety assessment, referrals, and follow-up for IPV was also a consistent prob-

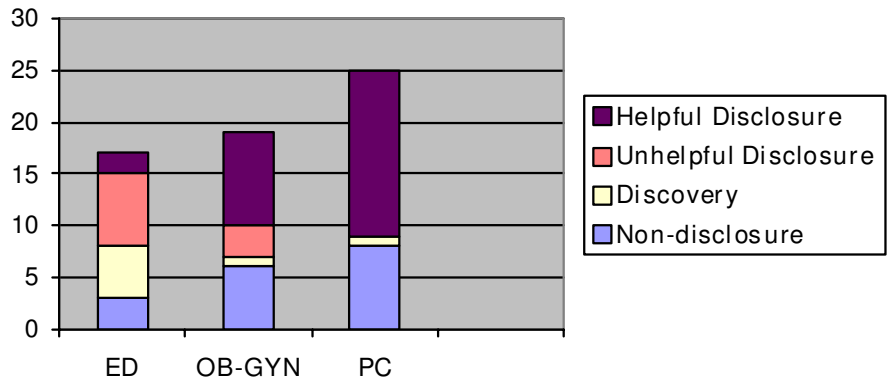


Figure 2
Encounter Attribute by Specialty.

lem in unhelpful disclosures. During her prenatal intake visit, for example, one participant disclosed ongoing violence, the name of her abusive partner, and his status as an undocumented foreigner. Not understanding what follow-up would occur, she became petrified that her husband would assault her for revealing his status. She subsequently switched to another prenatal care provider, where she lied about her home life. Several participants disclosed abuse but reported receiving little helpful advice from the clinician. The woman encouraged to contact the police above did not remember being offered contact information for safe houses in the area. Similarly, after treatment for acute injuries, one participant reported: "I don't recall ever getting information about a shelter... or an advocate speaking to me. Any of that."

Benefits of Disclosure: Making Changes, Improving Self-Esteem, Building Relationships

Eleven of the 25 beneficial disclosure experiences led directly to a change in the participant's circumstances, such as leaving the abusive spouse, entering a detoxification program or filing a police report. For example, one participant with newborn twins and a toddler reported that after disclosing her husband's assaults,

> "*I started off in a shelter in* [a distant town] *because I couldn't find one that would take all of us right away... My* [OB] *got on the phone with Social Service to try and get me all the help that I needed.*"

In others, changes resulted after a clinician worked with a participant over a period of time. One participant described the effect of the close relationship she developed with her midwife:

> "*She was real supportive through my pregnancy, and told me 'everything will be okay,' and I'll be a good mother. And I am a good mother. 'Cause it made me realize a lot of things... that I was thinkin', and I had my whole life to live, but now I could do better with myself, as well as with my daughter... I'm workin' on gettin' housin', takin' care of my schoolin', just bein' responsible.*"

Instead of an immediate end to the abuse, these patient-clinician encounters resulted in a shift in the participant's self-esteem, perception of the violent relationship, or awareness of alternatives, eventually empowering her to seek help for the abuse on her own. For example, clinicians' assurances that relationship violence was unacceptable resonated with several participants. One participant reported her primary care doctor's sympathetic insistence that the batterer's behavior was wrong set the stage for her to take action:

> "*She was like, 'No... no one who loves you will put their hands on you.' You know, it's not right. 'That's not real love.'...After* [he broke] *the wrist, I said, 'No more.'*"

After being treated in the ED for IPV-related injuries, another participant left with information on local safe houses that she later consulted when she was ready to leave her abuser.

Some of the beneficial disclosure experiences resulted in a more positive attitude toward health care in general, as in 5 instances where participants reported feeling a greater closeness with their clinicians despite no other change in their circumstances. Whether or not disclosure led to change, analysis revealed three common characteristics of provider behavior in beneficial disclosures: 1) explicit acknowledgement of the content of the disclosure (all cases), 2) demonstration of a caring attitude after disclosure (most cases) and 3) specific referral to other resources (some cases). For example, one participant said an ED clinician explicitly acknowledged her abuse and demonstrated concern:

> "*He said, well, 'I hear you're in a battered women's shelter. What's the deal? I take a special interest in domestic violence and what happens,' and he sat and talked to me. I felt comfortable in talking to him because he was showing this special interest in what was going on with me.*"

Also of note, in all but two beneficial disclosures the participant reported familiarity with the clinician. In primary care, these relationships involved getting to know the clinician through a variety of contacts both related and unrelated to the IPV. In OB/GYN, these relationships generally formed during prenatal care, or in the peri-partum period when the participant had daily contact with hospital clinicians. Such familiarity can also occur in the ED setting, as in one case where the participant accepted advice from a nurse who had treated her a few weeks earlier for IPV-related injuries. When the participant returned to the ED with more injuries, the nurse recognized her:

> "*And I started crying, and she's like, 'Two weeks ago you was here, now you're back here again today and it's for the same thing. Your face isn't all bruised up like it was two weeks ago, but you're hurtin'. What's goin' on?' I broke down and told her...She was like, 'Well, you don't need to be in a relationship like that.*"'

The participant acted on referrals and left her abusive partner as a result of this encounter.

Potential Benefits and Problems without Disclosure

The common thread to benefits and problems without verbal disclosure by the participant included explicit clini-

cian acknowledgement of potential abuse (or lack thereof). In particular, participants reported being upset by health care providers who they felt should have recognized IPV but did not acknowledge it. This, in turn, led to avoidance of healthcare. One participant reported that healthcare personnel failed to bring up IPV even after her husband yelled at her in the ED during two separate visits. She interpreted this lack of acknowledgement as an indication that clinicians did not care to get more involved. Another participant was particularly disappointed that her primary care clinician did not address the abuse with her, given that she had received counseling about it from his nursing staff: "He never gave me any type of indication...he didn't talk to me about it. That's why I left him...because he wasn't really direct with me."

Several participants reported benefit when the clinician spoke openly with the participant about IPV but did not insist upon disclosure. Furthermore, clinicians in these encounters used verbal and non-verbal cues to convey concern, and offered options for intervention while not forcing the participant to take action. The aftermath of acute injury was a particularly vulnerable time, as survivors were emotionally and physically exhausted as well as fearful of more injury.

> "*They asked me, 'How did it happen?' 'What happened to you?' 'Who did that?' I was in so much pain that I really didn't want to talk about it.*"

A critical component of beneficial non-disclosure experiences was consideration of the patient's safety, as in this ED visit:

> "*She realized that I had other bruises on me. I thought he might hear her and I was like, 'No. Let's just drop the conversation. Let's just get me stitched up.' My husband came in so there was no more talk about it. When I left, she called me apart, and she* [said]: *'you could call here in an emergency and we could get you some help.*"'

Another example included ED staff suggestion that a participant treated for acute injuries continue care in PC: "and they gave me a choice, 'would you rather go to your doctor and tell them what happened?"' As a result of that referral, she revealed the abuse to her primary care clinician.

Discussion

Narratives of intimate partner violence survivors reveal no actual harms occurred as a result of disclosure of abuse to health care clinicians. However, some negative disclosure experiences did impair subsequent interactions with the health care provider as well as increase emotional distress. The benefits included immediate changes (e.g. filing a restraining order), improvement in self-esteem to facilitate long term changes, and relationship building with health care clinicians. The setting of care appeared to influence these outcomes, impacted strongly by patient familiarity with the clinician.

This study reinforces insights from prior studies that asking about IPV in longitudinal care specialties offers the greatest opportunity for disclosure [20]. Indeed, participants valued clinicians who knew them well over time and were thus more likely to find disclosure in such settings beneficial. The benefits of disclosure reported here went beyond simply providing information, as might have been expected, but suggest an impact on patient self-worth and empowerment. This suggests that the relationship between clinician and patient can itself be a point of healing, and should reassure clinicians that extensive training in domestic violence or counseling is not as important as nurturing the relationship with a patient [24,25].

In all specialties, participants were more likely to disclose IPV and find disclosure beneficial if clinicians respectfully addressed the abuse, ensured participants' physical safety after an assault, assured participants of confidentiality regarding disclosed information, gave patient choices for action and demonstrated emotional support. Indeed, our study demonstrates that inquiry and discussion of IPV in the right setting can be a powerful tool for change.

Despite the increased potential to identify and refer a victim of IPV in the aftermath of an acute injury [26,27], participants in this study had mixed experiences with disclosure in the ED. Someone being treated for an acute injury as a direct result of IPV is likely to be in a highly charged state from the physical and emotional pain [28]. These women may feel particularly vulnerable and sensitive to any perceived failure of empathy on the part of the clinician. Furthermore, the probability of a beneficial disclosure in the ED may be lower with lack of familiarity with the clinician, a key element in many helpful disclosures. There may also be organizational barriers [29], such as lack of stretches of time to spend with any one patient while trying to manage an emergency department with multiple patients with differing levels of acuity. Rhodes's analysis of audio-taped encounters between physicians and IPV survivors confirms the difficulty exhibited by many clinicians' attempting to address this issue [21]. Thus, the ability to process and receive help related to IPV may be higher if it is done outside the context of emergency care. Treatments for acute injury related to IPV should also be viewed by clinicians as opportunities to educate and empower the patient, leaving her with options to exercise when she is ready. This may empower clinicians as well if they feel they have a task in helping the

patient rather than just uncovering a painful problem. Because ongoing relationships were more likely to lead to helpful disclosure experiences in this study, acute care providers should, with the patient's permission, inform her regular clinician of her visit [30].

IPV case-finding may satisfy the need for a quantifiable, appropriate quality improvement measure. However, measuring case finding alone may obscure whether the inquiry is occurring in an empowering and safe manner that benefits survivors. In settings such as the ED or even inpatient hospital care, where the risks of disclosure may be higher, other measures of quality could include surveys of patients at high risk for IPV to assess whether they received any education about resources or options for IPV. Future studies of intervention for IPV could consider measuring empowerment and trust around IPV disclosure in the health care setting. Outcome measures often determine the emphasis of clinical care [31,32]. If an organization such as the Joint Commission chose a process measure of IPV education and patient empowerment, it might spur clinical practice to change.

There are several limitations to this study. First, we were not always able to determine the exact nature of the visit or specialty. Furthermore, participants were not directly asked to compare their experiences; differences were gleaned from the stories they told. This is typical of qualitative research studies in which unexpected themes emerge from close examination of the data. Self-report is subject to recall bias, which may be particularly affected by any post-traumatic stress disorder related to abuse. The interviews occurred almost 10 years ago and clinician response might have improved since then, given the educational efforts with medical students and residents. However, this has not been demonstrated in more recent studies [21]. Finally, we only interviewed women who had used community resources and may not represent all IPV survivors.

Conclusion

Our results reveal that whether or not disclosure of abuse is achieved, clinician conversations with survivors about IPV have a powerful impact on both positive and negative outcomes. When these conversations occur in the context of a supportive relationship with that clinician, positive outcomes are more likely. Although these findings will need to be replicated in other settings, this study suggests a need to tailor interventions for women who experience IPV to the nature of the clinical specialty, particularly treatment of acute injury. Our findings indicate that it is not enough for health care providers to simply ask about abuse. Clinicians should aim for a therapeutic relationship with IPV survivors that does not demand disclosure or action, but instead empowers and educates the patient.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

JL designed the study. JL and TB conducted the interviews. All coauthors helped analyze data and reviewed the manuscript drafted by JL for important intellectual content. All coauthors approved the final draft.

References

1. Coker AL, Davis KE, Arias I, Desai S, Sanderson M, Brandt HM, Smith PH: **Physical and mental health effects of intimate partner violence for men and women.** *Am J Prev Med* 2002, **23(4):**260-268.
2. Coker AL, Smith PH, Bethea L, King MR, McKeown RE: **Physical health consequences of physical and psychological intimate partner violence.** *Arch Fam Med* 2000, **9(5):**451-457.
3. Liebschutz JM, Mulvey KP, Samet JH: **Victimization among substance-abusing women: worse health outcomes.** *Arch Intern Med* 1997, **157:**1093-1097.
4. Frayne SM, Skinner KM, Sullivan LM, Tripp TJ, Hankin CS, Kressin NR, Miller DR: **Medical profile of women Veterans Administration outpatients who report a history of sexual assault occurring while in the military.** *J Womens Health Gend Based Med* 1999, **8(6):**835-845.
5. Campbell J, Jones AS, Dienemann J, Kub J, Schollenberger J, O'Campo P, Gielen AC, Wynne C: **Intimate partner violence and physical health consequences.** *Arch Intern Med* 2002, **162(10):**1157-1163.
6. Lipsky S, Caetano R: **Impact of Intimate Partner Violence on Unmet Need for Mental Health Care: Results From the NSDUH.** *Psychiatr Serv* 2007, **58(6):**822-829.
7. Family Violence Prevention Fund: **National Consensus Guidelines on Identifying and Responding to Domestic Violence Victimization in Health Care Settings.** San Francisco CA, The Family Violence Prevention Fund; 2004.
8. Rhodes KV, Drum M, Anliker E, Frankel RM, Howes DS, Levinson W: **Lowering the threshold for discussions of domestic violence: a randomized controlled trial of computer screening.** *Arch Intern Med* 2006, **166(10):**1107-1114.
9. Friedman LS, Samet JH, Roberts MS, Hudlin M, Hans P: **Inquiry about victimization experiences.** *Arch Intern Med* 1992, **152:**1186-1190.
10. Feder GS, Hutson M, Ramsay J, Taket AR: **Women Exposed to Intimate Partner Violence.** *Arch Intern Med* 2006, **166:**22-37.
11. United States Preventive Services Task Force: **Screening for Family and Intimate Partner Violence: Recommendation Statement.** *Ann Intern Med* 2004, **140:**382-386.
12. Anglin D, Sachs C: **Preventive care in the emergency department: screening for domestic violence in the emergency department.** *Academic Emergency Medicine* 2003, **10(10):**1118-1127.
13. Thompson RS, Rivara FP, Thompson DC, Barlow WE, Sugg NK, Maiuro RD, Rubanowice DM: **Identification and management of domestic violence: a randomized trial.** *Am J Prev Med* 2000, **19(4):**253-263.
14. Plichta SB: **Intimate partner violence and physical health consequences: policy and practice implications.** *Journal of Interpersonal Violence* 2004, **19(11):**1296-1323.
15. Wathen CN, MacMillan HL: **Interventions for violence against women: scientific review.** *JAMA* 2003, **289(5):**589-600.
16. Rhodes KV, Lauderdale DS, He T, Howes DS, Levinson W: **"Between me and the computer": increased detection of intimate partner violence using a computer questionnaire.** *Ann Emerg Med* 2002, **40(5):**476-484.
17. MacMillan HL, Wathen CN, Jamieson E, Boyle M, McNutt LA, Worster A, Lent B, Webb M, McMaster Violence Against Women Research G: **Approaches to screening for intimate partner violence in health care settings: a randomized trial.** *JAMA* 2006, **296(5):**530-536.
18. McCloskey LA, Lichter E, Williams C, Gerber M, Wittenberg E, Ganz M: **Assessing intimate partner violence in health care settings leads to women's receipt of interventions and improved health.** *Public Health Reports* 2006, **121(4):**435-444.

19. McCord-Duncan EC, Floyd M, Kemp EC, Bailey B, Lang F: **Detecting Potential Intimate Partner Violence: Which Approach do Women Want?** *Fam Med* 2006, **38(6):**416-422.
20. McCloskey LA, Lichter E, Ganz ML, Williams CM, Gerber MR, Sege R, Stair T, Herbert B: **Intimate partner violence and patient screening across medical specialties.** *Academic Emergency Medicine* 2005, **12(8):**712-722.
21. Rhodes KV, Frankel RM, Levinthal N, Prenoveau E, Bailey J, Levinson W: **"You're not a victim of domestic violence, are you?" Provider patient communication about domestic violence.** *Ann Intern Med* 2007, **147(9):**620-627.
22. Battaglia T, Finley E, Liebschutz J: **Survivors of intimate partner violence speak out: Trust in the patient-provider relationship.** *JGIM* 2003, **18(8):**.
23. Glaser B, Strauss A: **The Discovery of Grounded Theory.** Chicago , Aldine; 1967.
24. Kelly U: **"What will happen if I tell you?" Battered Latina women's experiences of health care.** *Can J Nurs Res* 2006, **38(4):**78-95.
25. Bates L, Hancock L, Peterkin D: **"A little encouragement": health services and domestic violence.** *Int J Health Care Qual Assur Inc Leadersh Health Serv* 2001, **14(2-3):**49-56.
26. Abbott J, Johnson R, Koziol-McLain J, Lowenstein SR: **Domestic violence against women: Incidence and prevalence in an emergency department population.** *JAMA* 1995, **273(22):**1763-1767.
27. Ernst AA, Nick TG, Weiss SJ, Houry D, Mills T: **Domestic violence in an inner-city ED.** *Ann Emerg Med* 1997, **30(2):**190-197.
28. Yam M: **Seen but not heard: battered women's perceptions of the ED experience.** *Journal of Emergency Nursing* 2000, **26(5):**464-470.
29. Gerbert B, Johnston K, Caspers N, Bleecker T, Woods A, Rosenbaum A: **Experiences of battered women in health care settings: a qualitative study.** *Women Health* 1996, **24(3):**1-17.
30. Glass N, Dearwater S, Campbell J: **Intimate partner violence screening and intervention: data from eleven Pennsylvania and California community hospital emergency departments.** *Journal of Emergency Nursing* 2001, **27(2):**141-149.
31. Burke SM: **Continuous compliance in the environment of care.** *J Healthc Prot Manage* 2005, **21(2):**6-14.
32. Mobley D, Bodin S: **Joint Commission preparation in a large academic medical center: a framework for success.** *J Healthc Qual* 1998, **20(5):**8-14.

Pre-publication history

The pre-publication history for this paper can be accessed here:

http://www.biomedcentral.com/1471-2458/8/229/prepub

Limited Literacy and Psychiatric Disorders Among Users of an Urban Safety-Net Hospital's Mental Health Outpatient Clinic

Alisa Lincoln, PhD, MPH,†‡ Dennis Espejo, MPH,†‡ Peggy Johnson, MD,‡ Michael Paasche-Orlow, MD, MPH,§ Jeanne L. Speckman, MSc,‖ Terri L. Webber, MPH,† and Roberta F. White, PhD¶#*

Abstract: Little is known about the relationship between mental illness and literacy despite both being prevalent problems. We examine whether literacy varies by psychiatric diagnoses. Interviews and chart reviews ($N = 100$) were conducted in a behavioral health outpatient clinic. The relationships among sociodemographics, rapid estimate of adult literacy in medicine, measures of verbal and visual intellectual abilities, and psychiatric diagnoses were examined. The mean rapid estimate of adult literacy in medicine score was 55.9 which is equivalent to below an eighth grade literacy level. Psychotic disorder ($p = 0.03$) was associated with limited literacy, and substance abuse ($p = 0.003$) and PTSD ($p = 0.07$) were associated with higher literacy in bivariate analyses. These diagnoses were further examined in multivariate models. Limitations include the small sample size and the over-representation of people with high levels of education. Increasing our understanding of the relationships between health literacy and psychiatric disorders will help inform the development of appropriate psychiatric care and better outcomes.

Key Words: Psychiatric diagnoses, literacy, outpatient clinic.

(*J Nerv Ment Dis* 2008;196: 687–693)

*Health Science and Sociology Departments, Northeastern University, Boston, Massachusetts; †Social and Behavioral Sciences Department, Boston University School of Public Health, Boston, Massachusetts; ‡Division of Psychiatry, Boston University School of Medicine and Boston Medical Center, Boston, Massachusetts; §Section of General Internal Medicine, Department of Medicine, Boston University School of Medicine and Boston Medical Center, Boston, Massachusetts; ‖Section of General Internal Medicine (Health Services Research Unit), Department of Medicine, Boston University School of Medicine and Boston Medical Center, Boston, Massachusetts; ¶Department of Environmental Health, Boston University School of Public Health, Boston, Massachusetts; and #Department of Neurology, Boston University School of Medicine and Boston Medical Center, Boston, Massachusetts.

Supported by a grant from the Boston University School of Public Health.

Send reprint requests to Alisa K. Lincoln, PhD, MPH, Departments of Health Science and Sociology, Northeastern University, 316 Robinson Hall, 360 Huntington Avenue, Boston, MA 02115. E-mail: al.lincoln@neu.edu.

ISSN: 0022-3018/08/19609-0687
DOI: 10.1097/NMD.0b013e31817d0181

Limited literacy and mental illness are both common in the United States. One in 5 American adults suffers from a mental illness annually (U.S. Department of Health and Human Services, 1999), and according to the Institute of Medicine, 90 million people in the United States lack the literacy proficiency necessary to properly understand and act on health information (National Center for Education Statistics, 2005; Nielsen-Bohlman et al., 2004; Paasche-Orlow et al., 1999). This has often been referred to as limited health literacy and reflects both a patient's literacy skills and the complexity of the tasks required by the health care system (Baker, 2006). The National Assessment of Adult Literacy, conducted in 2003, found that the majority of adults (53%) had intermediate health literacy, 11% had proficient health literacy, 22% basic health literacy, and 14% had below basic health literacy (Kutner et al., 2006). Recent studies have also suggested that limited literacy and mental illness commonly cooccur. One study estimated that approximately 54% of patients in a busy urban psychiatric clinic had limited literacy skills (Currier et al., 2001). It has also been reported in a nationally representative sample that 75% of the people with a self-reported mental health problem had limited literacy skills (Kirsch et al., 1993). Strikingly, over 20% of these adults had completed high school (Kefalides, 1999). Although the development of literacy skills is an important goal of education, the level of education achieved does not accurately predict literacy skill (Miles and Davis, 1995).

There is a growing literature associating limited literacy with negative health outcomes for cancer, diabetes, asthma, and hypertension (Bennett et al., 1998; Davis et al., 1996; Paasche-Orlow et al., 1999; Williams et al., 1998). These outcomes include more frequent hospitalization (Baker et al., 1998, 2002) and worse prevention practices in people with diabetes (Rothman et al., 2004; Schillinger et al., 2004), asthma (Paasche-Orlow et al., 1999; Williams et al., 1998), cancer (Davis et al., 1996; Scott et al., 2002; Sharp et al., 2002), and other chronic illnesses (Kalichman and Rompa, 2000; Kalichman et al., 1999). It is also associated with higher mortality rates (Sudore et al., 2006; Wolf et al., 2006). In these studies, limited literacy is recognized as a barrier to effective services (Currier et al., 2001) and to appropriate service utilization (Baker et al., 1998).

The recent Agency for Healthcare Research and Quality evidence review, "Literacy and Health Outcomes: Evidence Report/Technology Assessment," cited 5 studies that evaluate the association between a marker of health literacy and a marker of mental illness (Berkman et al., 2004). Four of these studies reported statistically significant associations between limited literacy and higher prevalence of depression, but not all of the associations remained significant in adjusted analyses. For example, Gazmarian et al. (2000) found that 13% of new Medicare recipients met criteria for depression on the Geriatric Depression Scale. Individuals with limited health literacy were 3 times more likely to be depressed. However, after controlling for demographics, social support, health behavior, and health status, health literacy did not remain an independent risk factor for depressive symptoms. Recently, in the first longitudinal analysis to address the issue of literacy and mental health symptoms, limited literacy was linked to a higher level of depressive symptoms in patients with addiction (Lincoln et al., 2006).

Literacy impacts health in many ways; however, limited literacy presents unique challenges in mental health care. Standard psychiatric evaluation does not include assessment of literacy. Current best-practice in mental health care includes several treatment modalities which assume adequate literacy such as journaling in cognitive behavioral therapy. People struggling with limited literacy may not benefit fully from these treatments or may be excluded altogether if literacy is not addressed.

In this article, we present the results of a study conducted in a public, urban, mental health outpatient clinic. The study was initiated as a first step to understand the complex mechanisms of association between mental illness and literacy and was designed to examine the full range of psychiatric disorders. Thus, this study expands on the previous literature, which is primarily focused on depressive symptoms and depression. The study assesses the prevalence of limited literacy among a convenient sample of 100 people seeking services at the clinic and examines the relationship between literacy level and a wide range of psychiatric diagnoses. The study tested the hypothesis that the prevalence of limited literacy will vary with psychiatric diagnosis.

METHODS

Data Collection

Data were collected over a period of 8 months in 2004/2005 in the behavioral outpatient clinic of a busy, public, urban safety-net hospital. Two interviewers recruited participants and conducted interviews on varying days of the week and times of the month to account for weekly or monthly service utilization patterns. Patients who had appointments at the Behavioral Outpatient Clinic on study days were approached by 1 of the interviewers while sitting in the waiting room. In the pilot phase (12 interviews), potential participants were told that the purpose of the study was to explore the relationship between mental health and literacy. However, the patients who agreed to participate in the pilot had disproportionately high literacy scores. It may be that the stigma of limited literacy deterred participants with limited literacy from participating in a study presented as focusing on literacy. As a consequence, recruitment procedures were revised.

In the study reported here, patients were asked if they would like to participate in a brief study of education and health. After completing the description of the study to the participants, written informed consent was obtained. Sixty-seven percent of patients approached in the waiting room chose not to participate in the study. The most frequently cited reason was that they did not have time before their appointment. An incentive of $10 was paid upon completion of the interview. All participants were age 18 and over and the only exclusion criterion was not being able to speak English. The study was approved by the Boston University Medical Center Institutional Review Board.

The demographic characteristics of the study sample in large part reflect the population served in the clinic and other public urban settings (Table 1). The sample was 61% men and 39% women, 29% White, 48% African American, 11% Hispanic, and 12% other race/ethnicity. The mean age of the

TABLE 1. Demographics (N = 100)

	N
Sex	
Male	61
Female	39
Race	
Caucasian	29
African American	48
Hispanic	11
Other	12
Age	
18–29 yrs	13
30–39 yrs	28
40–49 yrs	36
50–59 yrs	18
60+ yrs	5
Education	
<High school	36
High school of GED	39
>High school	25
Primary language	
English	92
Other language	7
Any axis I diagnosis of chart review	
Depression	12
Psychotic disorder	23
Bipolar	55
Anxiety/panic disorder	14
Substance abuse	38
ADHD	5
PTSD	19
No diagnosis	3
Mean REALM score (*SD*)	55.88 (13.89); range: 3–66
Mean WAIS score (*SD*)	16.6 (6.25); range: 0–30
Mean Ravens score (*SD*)	15.6 (4.82); range: 3–24

sample was 41.6 years (standard deviation 10.7), which did not differ from the mean age of the population seen in the clinic (41.3 years). However, the study sample was more highly educated than the population seen in the clinic, with 25% having more than a high school education, as compared with 12% of the clinic population.

Data were collected through structured interview followed by a structured medical record review. Interview data included participants' demographic and socioeconomic information, including age, gender, race, ethnicity, primary language, education, and work status. Literacy was assessed using the rapid estimate of adult literacy in medicine (REALM) (Davis et al., 1993). REALM scores were used to determine literacy level. Higher literacy was defined as a REALM score of 61 to 66, whereas limited literacy was defined as a score of 0 to 60 on the REALM. The REALM describes scores of below 60 as roughly equivalent to below eighth grade level. In addition, the authors noted that subjects who score below 60 on the REALM "will struggle with most patient education materials." Subjects with REALM scores below 45 "will need low literacy materials; may not be able to read prescription labels," and subjects with REALM scores below 19 "will not be able to read most low literacy materials; will need repeated oral instructions, materials composed primarily of illustrations, or audio or video tapes." (Davis et al., 1993). To dissociate reading skills from general cognitive abilities, 2 measures were used. The Wechsler Adult Intelligence Scale-III (WAIS-III) Similarities subtest (Wechsler, 1997) was used to assess verbal intelligence. This task is generally acceptable to patient participants, and it assesses reasoning-based cognitive skills rather than classroom-based information and abilities. Total correct raw scores using WAIS-III scoring criteria were used as outcomes in data analysis. Similarly, the Raven Standard Progressive Matrices Test (Raven, 1958) was employed to test reasoning skills using visual information. Only the first 2 of the 5 series that comprise the test were administered because of time constraints, and the outcome was the number of correct responses.

The choice of instruments and measures was guided by the theoretical foundation that verbal skills/verbally based problem solving (as measured by the WAIS Similarities subtest) and visuospatial skills (as measured by the Raven's Progressive Matrices) and education are all components of reading abilities. Both the Similarities and the Raven can be performed well in people who have disassociations between reading and intellect. Both measures were pilot tested for feasibility with this population. Because, only part of the Raven was administered, these data were not appropriate for the final analyses. However, we present these pilot data here as this preliminary examination suggests this test is a feasible measure for this population. Data from the Similarities were complete, and thus, these data are presented in the final analyses.

In a second stage, participants' medical records were reviewed to gather additional data on psychiatric diagnoses (axis I and II) and medical diseases (axis III). Axis I diagnoses were coded as primary, secondary, and other; and "rule out" diagnoses and history of the disorders were noted. In addition, the number of visits to the outpatient clinic in the previous year, family/social network involvement, and data on literacy or other education related issues described in the patient's medical record were recorded.

Axis I diagnosis codes were grouped into 8 categories: depression, psychotic disorders, bipolar disorders, anxiety and panic disorders (other than PTSD), substance abuse disorders, posttraumatic stress disorder, attention deficit hyperactivity disorder, and other disorders. Participants with multiple diagnoses were considered to have diagnoses in multiple categories. "Rule out" diagnoses were not considered a diagnosis and were not categorized. Two individuals, for whom no axis I diagnosis was recorded, were grouped into a "no diagnosis" category.

Information on race and ethnicity was coded as: White, Black, non-Hispanic, Hispanic, and Other. Education was coded as less than high school, graduated high school or received a GED, more than high school, and unknown.

All analyses were performed with SAS 8.02 (SAS Institute, Cary, NC). Bivariate statistics were computed by chi square (for dichotomous outcomes, i.e., limited literacy) or Wilcoxon rank sum test (for continuous outcomes, i.e., REALM score). Multivariate models were run using logistic models. All tests were performed as two tailed with a 95% confidence interval.

RESULTS

Twelve percent of the participants had diagnoses of depression, 24% psychotic disorders, 55% bipolar disorder, 14% anxiety/panic, 38% substance abuse, 19% PTSD, 5% attention deficit hyperactivity disorder, and 2% of participants had no diagnosis noted in the medical record (Table 1).

In addition, further analyses of the diagnostic data demonstrated a high level of comorbidity among psychiatric diagnoses with 31% of the sample having 2 diagnoses and 21% having 3 or more diagnoses. The primary axis I diagnosis with the highest number of comorbid disorders was bipolar disorder. Fifty people had a primary diagnosis of bipolar disorder. Among those 50 participants, 2 had an additional diagnosis of a psychotic disorder, 7 had an anxiety disorder, 28 had substance abuse, and 16 had other disorders noted.

The mean REALM score for the sample was 55.9, which indicates an average reading level of seventh to eighth grade. The mean raw score on the WAIS-III Similarities subtest was 16.6 which is about 1 standard deviation below the average score for the sample mean age (41.6 years). The mean Raven score was 15.6 for the first 2 series of stimuli. The Raven was used to test the feasibility of the use of this measure with this population, however, not used in further analyses because the entire test was not administered. It was found to be feasible and acceptable to participants and the Raven was highly correlated with both Similarities (0.502) and REALM (0.420). Finally, Similarities raw scores and REALM scores were highly correlated (0.420), as expected.

In unadjusted analyses, limited literacy was significantly associated with lower levels of formal education and diagnosis of a psychotic disorder. Higher literacy was associated with diagnoses of substance abuse disorder and PTSD (Table 2).

TABLE 2. Bivariate Analysis With REALM Scores (N = 100)

	Limited Literacy			
Variable	**N**	**%**	χ^2	**Sig. Level**
Education			15.4	<0.01
Less than high school	36	61		
High school or GED	39	51		
More than high school	25	12		
Psychotic disorder	23	65	4.9	0.03
No psychotic disorder	77	39		
Substance abuse	38	26	8.65	<0.01
No substance abuse	62	56		
PTSD	19	26	3.31	0.07
No PTSD	81	49		
Anxiety disorder	14	29 (4)	1.78	0.18
No anxiety disorder	86	48 (41)		
Depression	12	50 (6)	0.14	0.71
No depression	88	44 (39)		
Bipolar	55	40 (22)	1.23	0.27
No bipolar	45	51 (23)		

Initial analyses demonstrated high correlations between age and both Similarities and REALM. To address this, several multivariate models were created for predicting limited literacy. In each set, we used the diagnoses which were statistically significant predictors of literacy outcomes in bivariate analyses: psychotic disorder, substance abuse disorder, and PTSD. In the first set of models (Table 3), presence of each diagnosis and then the 3 dichotomous diagnoses variables in combination to capture the high level of comorbid disorders in this sample were used to predict limited literacy. Each of these models was adjusted for race/ethnicity and age. Education was not included in these models. Although education was correlated with REALM score, it was not correlated with any of the independent variables. In these adjusted models, people with a psychotic disorder were 2.7 times more likely ($p = 0.06$) as those without a psychotic disorder to meet criteria for limited literacy. People with substance abuse disorder were one-third less likely [odds ratio (OR) = 0.29, $p < 0.01$] to have limited literacy (or 3 times more likely to have higher literacy) than people without substance abuse disorders, and those with PTSD were less than half as likely (OR = 0.42, $p < 0.15$) to have limited literacy as those without. These relationships remained when the diagnoses were entered into the final model in combination.

In a second set of models (Table 4) adjusted for race/ethnicity, age, and the WAIS-III Similarities raw score, patients with psychotic disorders were more likely to meet criteria for limited literacy (OR = 2.4, $p < 0.14$) than people without psychotic disorders, and people with substance abuse and PTSD were less likely to have limited literacy than people without these disorders (OR = 0.29, $p < 0.02$ and 0.39, $p < 0.14$, respectively). Again, in the final model (model 4) when the diagnoses are entered into the model in combination, the relationships persist. Thus, when additionally adjusting for verbal reasoning abilities, the findings are consistent with Table 3.

In a final analytic step, analyses were conducted on the subset of participants under the age of 50, because cognitive function and age may be particularly highly correlated among participants over the age of 50, and this was born out in bivariate analysis. In these models (Table 5), people with psychotic disorder were almost 5 times as likely as people without psychotic disorder to have limited literacy ($p < 0.05$). People with substance abuse disorder were one-fifth as likely as people without substance abuse disorder to have limited literacy (or 5 times as likely as people without substance abuse disorder to have higher literacy, $p < 0.01$). Finally, in this model, PTSD does not significantly predict literacy, although results are similar in direction to the previous model (OR = 0.39, $p < 0.13$). These models again show that the inclusion of each of the diagnostic variables in combination does not change the relationships between diagnoses and literacy.

Finally, given the high level of psychiatric comorbidity among the sample, analyses were conducted examining the number of diagnoses and literacy. The number of diagnoses did not predict literacy in similar models, and thus these data are not presented here.

TABLE 3. Logistic Regression Analysis of Diagnoses, Race/Ethnicity, and Age on Limited Literacy (N = 100)[a]

	Model 1		Model 2		Model 3		Model 4	
Variable	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**
Psychotic disorder	2.67*	(0.95, 7.44)					1.75	(0.58, 5.28)
Substance abuse disorder			0.29***	(0.11, 0.72)			0.43**	(0.13, .90)
PTSD					0.42	(0.13, 1.36)	0.53	(0.16, 1.80)
Race								
African American	1.55	(0.56, 4.30)	1.57	(0.56, 4.41)	1.77	(0.65, 4.82)	1.38	(0.48, 3.98)
Hispanic	3.39	(0.75, 15.24)	3.98*	(0.84, 18.84)	3.24	(0.72, 14.59)	3.58	(0.75, 16.98)
Other race	2.63	(0.60, 11.65)	2.82	(0.65, 12.31)	3.63*	(0.85, 15.47)	2.42	(0.53, 11.06)
Age	1.03	(0.98, 1.07)	1.03	(0.98, 1.07)	1.03	(0.98, 1.07)	1.02	(0.98, 1.07)

[a]Each model, except number 4, includes only 1 diagnosis. The reference group for race is white.
*$p < 0.10$.
**$p < 0.05$.
***$p < 0.01$.

TABLE 4. Logistic Regression Analysis of Diagnosis, Race/Ethnicity and WAIS III Similarities on Limited Literacy (N = 100)[a]

	Model 1		Model 2		Model 3		Model 4	
Variable	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**
Psychotic disorder	2.4**	(0.87, 7.44)					1.47	(0.42, 5.21)
Substance abuse disorder			0.29***	(0.11, 0.77)			0.31***	(0.11, 0.89)
PTSD					0.39	(0.11, 1.38)	0.41	(0.11, 1.54)
Race								
African American	1.14	(0.38, 3.43)	1.09	(0.36, 3.36)	1.18	(0.39, 3.57)	0.89	(0.28, 2.87)
Hispanic	1.52	(0.28, 8.19)	1.77	(0.31, 10.15)	1.36	(0.26, 7.20)	1.45	(0.25, 8.38)
Other race	2.20	(0.44, 10.96)	2.13	(0.45, 10.02)	2.69	(0.55, 13.08)	1.89	(0.37, 9.64)
WAIS III similarities	0.85****	(0.77, 0.93)	0.85****	(0.77, 0.93)	0.85****	(0.77, 0.93)	0.84***	(0.76, .93)
Age	1.03	(0.99, 1.08)	1.04	(0.99, 1.09)	1.03	(0.98, 1.08)	1.03	(0.98, 1.08)

[a]Each model, except number 4, includes only one diagnosis. The reference group for race is white.
*$p < 0.10$.
**$p < 0.08$.
***$p < 0.05$.
****$p < 0.01$.

TABLE 5. Logistic Regression Analysis of Diagnosis, Race/Ethnicity, Age and WAIS III Similarities on Limited Literacy for Participants Under Age of 50 (N = 77)[a]

	Model 1		Model 2		Model 3		Model 4	
Variable	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**	**OR**	**95% CI**
Psychotic disorder	4.73***	(1.02, 21.93)					3.50	(0.55, 16.65)
Substance abuse disorder			0.20****	(0.07, 0.63)			0.23***	(0.07, 0.74)
PTSD					0.39	(0.11, 1.31)	0.42	(0.11, 1.63)
Race								
African American	0.81	(0.22, 3.01)	0.95	(0.25, 3.55)	0.98	(0.28, 3.56)	0.63	(0.15, 2.62)
Hispanic	2.08	(0.38, 11.48)	2.77	(0.45, 16.91)	1.73	(0.32, 9.28)	2.44	(0.39, 15.07)
Other race	2.31	(0.41, 12.92)	2.43	(0.44, 13.37)	2.95	(0.54, 16.20)	1.87	(0.32, 10.98)
WAIS III similarities	0.93	(0.84, 1.02)	0.92*	(0.83, 1.01)	0.91***	(0.83, 1.00)	0.93	(0.83, 1.04)
Age	1.06*	(0.99, 1.14)	1.07***	(1.00, 1.15)	1.06	(0.99, 1.13)	1.00*	(1.00, 1.15)

[a]Each model, except number 4, includes only one diagnosis. The reference group for race is white.
*$p < 0.10$.
**$p < 0.08$.
***$p < 0.05$.
****$p < 0.01$.

DISCUSSION

In this article, we present the results of a study designed to assess levels of literacy among people seeking services for the full range of psychiatric disorders in a busy public urban safety-net hospital behavioral health clinic. Although there is increasing attention to the prevalence and importance of limited literacy and a variety of health outcomes, less attention has been paid in behavioral health. Few studies have examined literacy levels among patients with the full complement of psychiatric disorders, including the psychotic disorders, affective disorders, anxiety disorders, and substance abuse, and these have been limited by sample size and a lack of diagnostic data.

In these data, REALM score was highly associated with level of education and diagnosis of a psychotic disorder, a substance abuse disorder or PTSD in bivariate analyses. No bivariate association was found between anxiety disorders, depression or bipolar disorder, and REALM. In models using larger numbers of adjustments, psychotic disorders, substance abuse disorders, and PTSD all predicted limited literacy when adjusted for race/ethnicity, age and WAIS-III Similarities. However, the diagnoses seem to be related to literacy levels in different ways. In the adjusted models, the presence of a psychotic disorder was associated with increased likelihood of limited literacy by almost 3-fold. Importantly, the presence of either a substance abuse disorder or a diagnosis of PTSD was associated with a decreased likelihood of limited literacy. The relationships were similar, but the effect sizes are even greater when examined in the subsample of people under the age of 50. This apparent incongruity in the direction of the impact of psychiatric disorders on literacy levels has not been explored elsewhere. Certainly the psychotic disorders, primarily schizophrenia, are likely to interrupt the educational trajectory of people based on the average age of onset of symptoms. The finding of an increased likelihood of limited literacy among patients with psychotic disorders may reflect

this. The small sample size did not permit full exploration of this possibility, and further work must be done to explore this.

The study also demonstrated the feasibility of employing WAIS-III Similarities and the Raven as measures of verbal and visuospatial intelligence independent of education and literacy. These measures are important in future studies to disentangle the effects of cognitive limitations, education deprivation, and reading ability on literacy.

This study had several important limitations. First, the sample was small, and thus, future work is required to replicate and validate the findings. The sample size also has limited capacity to examine the significance of the predictors with each of the diagnostic variables in the final model. The relationships between both PTSD and psychotic disorders, and literacy do not remain statistically significant in models with each of the 3 diagnostic categories. More work is needed to determine whether this reflects unique relationships between diagnoses and literacy among patients with multiple disorders or is a reflection of our limited sample size. Secondly, patients' psychiatric diagnoses were determined by medical records instead of independent research assessments. This was driven by limited resources and the small size of the study, but future work should include full research diagnostics. The high proportion of patients' with bipolar disorder may be a reflection of diagnostic practice with patients who have high levels of comorbid substance abuse, in this setting. In addition, literacy levels might actually impact the diagnostic decisions made and thus noted in the medical record. Third, this is not a random sample, but a sample of convenience. Recruiting subjects while waiting for appointments resulted in a high refusal rate, and based on our pilot work, it is clear that there may be some bias in who chooses to speak with a researcher about education and health. However, we are confident that the sample is generally reflective of the people seeking care in the clinic, based on the frequency and types of diagnoses in the sample, and its gender, age, and racial make-up. The sample does however seem more highly educated than the general population seeking care in this setting. Finally, the results of this study may not be generalizable to people seeking services in other behavioral health settings or in less urban areas.

CONCLUSIONS

These data, some of the first to examine literacy across a full-range of psychiatric diagnoses and comorbid diagnoses, demonstrate that the relationships between limited literacy and psychiatric diagnosis are complex. Although patients with psychotic disorders were more likely to have limited literacy, the presence of PTSD or a substance abuse disorder was associated with a lower likelihood of limited literacy in this sample.

Future work should address the role of limited literacy among patients with all behavioral health needs, rather than focusing exclusively on patients with depression as has been the emphasis of the literature previously. In addition, future efforts would benefit from the use of a structured diagnostic interview to ascertain diagnoses. Because several best-practice psychotherapeutic treatments require adequate literacy, these data suggest that clinicians and providers should increase their attention to literacy levels and examine their strategies for patient assessment and education. In addition, successful rehabilitation approaches should include thorough assessments and consideration of literacy.

REFERENCES

Baker DW (2006) The meaning and the measure of health literacy. *J Gen Intern Med*. 8:878–883.

Baker DW, Gazmararian JA, Williams MV, Scott T, Parker RM, Green D, Ren J, Peel J (2002) Functional health literacy and the risk of hospital admission among medicare managed care enrollees. *Am J Public Health*. 92:1278–1283.

Baker DW, Parker RM, Williams MV, Clark WS (1998) Health literacy and the risk of hospital admission. *J Gen Intern Med*. 13:791–798.

Bennett CL, Ferreira MR, Davis TC, Kaplan J, Weinberger M, Kuzel T, Seday MA, Sartor O (1998) Relation between literacy, race and stage of presentation among low-income patients with prostate cancer. *J Clin Oncol*. 16:3101–3104.

Berkman ND, DeWalt DA, Pignone MP, Sheridan SL, Lohr KN, Lux L, Sutton SF, Swinson T, Bonito AJ (2004) Literacy and Health Outcomes. Evidence Report/Technology Assessment No. 87 (Prepared by RTI International-University of North Carolina Evidence-based Practice Center under Contract No. 290-02-0016). AHRQ Publication No. 04-E007-2. Rockville (MD): Agency for Healthcare Research and Quality.

Currier GW, Sitzman R, Trenton A (2001) Literacy in the psychiatric emergency service. *J Nerv Ment Dis*. 189:56–58.

Davis TC, Arnold C, Berkel HJ, Nandy I, Jackson RH, Glass J (1996) Knowledge and attitude on screening mammography among low-literate, low-income women. *Cancer*. 78:1912–1920.

Davis TC, Long SW, Jackson RH, Mayeaux EJ, George RB, Murphy PW, Crouch MA (1993) Rapid estimate of adult literacy in medicine: A shortened screening instrument. *Fam Med*. 25:391–395.

Gazmararian J, Baker D, Parker R, Blazer DG (2000) A multivariate analysis of factors associated with depression: Evaluating the role of health literacy as a potential contributor. *Arch Intern Med*. 160:3307–3314.

Kalichman SC, Ramachandran B, Catz S (1999) Adherence to combination antiretroviral therapies in HIV patients of low health literacy. *J Gen Intern Med*. 14:267–273.

Kalichman SC, Rompa D (2000) Functional health literacy is associated with health status and health-related knowledge in people living with HIV-AIDS. *J Acquir Immune Defic Syndr*. 25:337–344.

Kefalides PT (1999) Illiteracy: The silent barrier to health care. *Ann Intern Med*. 130:333–336.

Kirsch I, Jungeblut A, Jenkins L, Kolstad A (1993) *Adult Literacy in America: A First Look at the Findings of the National Adult Literacy Survey*. Washington (DC): National Center for Education Statistics.

Kutner M, Greenberg E, Jin Y, Paulson C (2006) *The Health Literacy of America's Adults: Results From the 2003 National Assessment of Adult Literacy* (*NCES Publication No.* 2006483). Washington (DC): National Center for Education Statistics.

Lincoln A, Paasche-Orlow MK, Cheng DM (2006) Impact of health literacy on depressive symptoms and mental health related quality of life among adults with addiction. *J Gen Intern Med*. 21:818–822.

Miles S, Davis T (1995) Patients who can't read: Implications for the health care system. *J Am Med Assoc*. 274:1719–1720.

National Center for Education Statistics (2005) National Assessment of Adult Literacy. A First Look at the Literacy of America's Adults in the 21st Century (NCES Publication No. 20006470). Washington (DC): National Center for Education Statistics.

Nielsen-Bohlman LT, Panzer AM, Kindig DA (Eds) (2004) Health Literacy: A Prescription to End Confusion. Committee on Health Literacy, Board on Neuroscience and Behavioral Health. Institute of Medicine. Washington (DC): National Academies Press.

Paasche-Orlow MK, Parker RM, Gazmararian JA, Nielsen-Bohlman LT, Rudd RR (1999) The prevalence of limited health literacy. *J Gen Intern Med*. 20:175–184.

Raven JC (1958). Standard Progressive Matrices. Cambridge, England: HK Lewis.

Rothman RL, DeWalt DA, Malone R, Bryant B, Shintani A, Crigler B,

Weingberger M, Pignone M (2004) Influence of patient literacy on the effectiveness of a primary care-based diabetes disease management program. *J Am Med Assoc*. 292:1711–1716.

Schillinger D, Grumbach K, Piette J, Wang F, Osmond D, Daher C, Palacios J, Sullivan GD, Bindman AB (2004) Association of health literacy with diabetes outcomes. *J Am Med Assoc*. 288:475–482.

Scott TL, Gazmararian JA, Williams MV, Baker DW (2002) Health literacy and preventive health care use among medicare enrollees in a managed care organization. *Med Care*. 40:395–404.

Sharp LK, Zurawski JM, Roland PY, O'Toole C, Hines J (2002) Health literacy, cervical cancer risk factors and distress in low-income African-American women seeking colposcopy. *Ethn Dis*. 12:541–546.

Sudore RL, Yaffe K, Satterfield S, Harris TB, Mehta KM, Simonsick EM, Newman AB, Rosano C, Rooks R, Rubin SM, Ayonahon HN, Schillinger D (2006) Limited literacy and mortality in the elderly: the health, aging and body composition study. *J Gen Intern Med*. 21:806–812.

U.S. Department of Health and Human Services (1999) *Mental Health: Report of the Surgeon General.* Rockville (MD): National Institutes of Health.

Wechsler D (1997) Wechsler Adult Intelligence Scale-III (WAISIII), San Antonio, TX: Psychological Corp.

Williams MV, Baker DW, Honig EG, Lee TM, Nowlan A (1998) Inadequate literacy is a barrier to asthma knowledge and self-care. *Chest*. 114:1008–1015.

Wolf MS, Feinglass JM, Carrion V (2006) Literacy and mortality among medicare enrollees. *J Gen Intern Med*. 21:81.

Regular Article

The BMC ACCESS Project: The Development of a Medically Enhanced Safe Haven Shelter

Alisa Lincoln, PhD, MPH
Peggy Johnson, MD
Dennis Espejo, MPH
Sara Plachta-Elliott, MA
Peggy Lester, Ed, MPA
Christopher Shanahan, MD, MPH
Susan Abbott, MED
Howard Cabral, PhD
Amber Jamanka, MPH
Jonathan Delman, JD, MPH
Patty Kenny, APRN, BC

Abstract

This paper describes the development and implementation of the Boston Medical Center (BMC) Advanced Clinical Capacity for Engagement, Safety, and Services Project. In October 2002, the

Address correspondence to Alisa Lincoln, PhD, MPH, Department of Health Sciences and Sociology, Northeastern University, 316 Robinson Hall, 360 Huntington Ave., Boston, MA 02115, USA. Email: al.lincoln@neu.edu.

Peggy Johnson, MD, Department of Psychiatry, Boston University School of Medicine, Boston, MA, USA.

Dennis Espejo, MPH, Department of Psychiatry, Boston University School of Medicine, Boston, MA, USA.

Dennis Espejo, MPH, Department of Social and Behavioral Sciences, Boston University School of Public Health, Boston, MA, USA.

Christopher Shanahan, MD, MPH, Section of General Internal Medicine, Boston University School of Medicine, Boston, MA, USA.

Susan Abbott, MED, Vinfen Corporation, Cambridge, MA, USA.

Peggy Lester, Ed, MPA, Metro Boston Area Office, Massachusetts Department of Mental Health, Boston, MA, USA.

Patty Kenny, APRN, BC, Metro Boston Area Office, Massachusetts Department of Mental Health, Boston, MA, USA.

Sara Plachta-Elliott, MA, Consumer Quality Initiatives, Roxbury, MA, USA.

Jonathan Delman, JD, MPH, Consumer Quality Initiatives, Roxbury, MA, USA.

Amber Jamanka, MPH, Data Coordinating Center, Boston University School of Public Health, Boston, MA, USA.

Howard Cabral, PhD, Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA

Present address: Alisa Lincoln, PhD, MPH, Department of Health Sciences and Sociology, Northeastern University, Boston, MA, USA.

Journal of Behavioral Health Services & Research, 2008.

BMC Division of Psychiatry became the first such entity to open a Safe Haven shelter for people who are chronically homeless, struggling with severe mental illness, and actively substance abusing. The low-demand Safe Haven model targets the most difficult to reach population and serves as a "portal of entry" to the mental health and addiction service systems. In this paper, the process by which this blended funded, multi-level collaboration, consisting of a medical center, state, city, local, and community-based consumer organizations, was created and is maintained, as well as the clinical model of care is described. Lessons learned from creating the Safe Haven Shelter and the development and implementation of the consumer-informed evaluation are discussed as well as implications for future work with this population.

Introduction

People who are chronically homeless and struggling with both severe mental illness and substance abuse remain among the most difficult to engage in housing and treatment. In 2002, the Division of Psychiatry at Boston Medical Center (BMC) developed a partnership with state, local, and community-based organizations to form the Advanced Clinical Capacity for Engagement, Safety and Services (ACCESS) Project. The BMC ACCESS Project filled a major gap, identified by the Massachusetts Department of Mental Health (DMH), in the existing adult services programming for adults dually diagnosed with substance abuse and mental illness who are chronically homeless in the city of Boston. The project designed and opened a clinically enhanced Safe Haven shelter to engage persons who are homeless in accessing integrated services such as mental health treatment, substance abuse treatment, and primary care and ultimately to promote residential stability. In addition, a consumer-informed evaluation of the project was conducted. This shelter is Boston's first and only Safe Haven shelter for men and women who are chronically homeless, actively using substances, and living with a severe mental illness. In addition, it may be the only Safe Haven shelter operating within a division of psychiatry in an academic medical center. In this paper, the development and implementation of this unique program model as well as the design of the consumer-informed evaluation is described. Results of the evaluation are forthcoming. Important lessons learned from this project and implications for future efforts to best meet the needs of people who are chronically homeless, actively substance using, and who have a severe mental illness are discussed.

Development of the BMC ACCESS Project

In Boston, Massachusetts from 1992 to 2002, the number of people who were homeless increased from 4,411 to 6,210 or by 41%.[1] Approximately one third of people who are homeless have a severe mental illness, and about half are co-morbid for alcohol and substance abuse disorders.[2–5] Homeless individuals have three to four times higher rates of death than the general population due to the heavy burden of disease and disadvantaged status.[6] Homeless individuals with mental illness are also characterized by poor physical health, past traumas, long-term poverty, social isolation, lack of vocational skills, and stigma associated with contact with the criminal justice system.[7]

The Division of Psychiatry at BMC initially partnered with the Metro Boston Area of the DMH in an effort to examine existing services for chronically homeless individuals dually diagnosed with substance abuse and mental illness in Boston. For many years, advocacy groups working with people who are chronically homeless in Boston had called for an increase in the number of Safe Haven beds in the city. Only six Safe Haven beds existed in Boston, and these were only available to women, had a long waiting list, and had been at maximum occupancy since opening. The Safe Haven model was created (Title IV, Subtitle D of the McKinney Act) to engage and house people

with mental illness and substance abuse issues who have been unwilling or unable to participate in supportive housing services. This transitional housing model was developed to address the special needs of this "resistant to treatment" population and reduce housing barriers unique to this group. Safe Havens and Housing First programs are similar in their focus on providing housing and not requiring the individual to be "housing ready" which often includes being sober or mentally stable. A distinctive difference is that Housing First programs typically provide their participants with direct links to permanent mainstream housing, while Safe Havens are viewed as transitional housing with the goal of transitioning to more permanent housing when the resident feels they are ready to move. Currently, there is a lack of data on the development of Safe Havens for the dually diagnosed population and particularly the impact of integrated mental health, substance abuse, and primary-care services.

A recent national survey of Safe Haven programs, conducted by the Ward Foundation, highlights the diversity that has developed with the implementation of the Safe Haven model.[8] Seventy-nine of the 118 programs identified participated in the survey. The average capacity of Safe Havens is 16 residents, with 63.3% offering residents a private room. Approximately 72.1% do not impose a limit on length of stay, and the average length of stay was 262.4 days. Admissions criteria varied, but the majority of the programs require a diagnosis of a severe and persistent mental illness (89.9%). Interestingly, 6.5% of the Safe Havens surveyed require residents to remain clean and sober for admittance into the program. Most are staffed 24 h/day (88.6%), and the majority of the Safe Havens were located in a mixed-housing neighborhood. While there were differences in the components of the program in terms of admissions criteria, physical space, and philosophies, the report concluded that Safe Havens are effective in engaging and retaining the residents, as over half of the residents were successfully housed in some type of permanent housing.

Following the initial examination of existing services for this population, the BMC Division of Psychiatry and the Metro Boston Area of the Massachusetts DMH developed a proposal in response to the Substance Abuse and Mental Health Services Administration (SAMHSA) Treatment for the Homeless initiative. In doing so, it became clear that the traditional Safe Haven model would benefit from an enhancement of services, and a model of a "medically enhanced" Safe Haven was developed. This required partnering with a diverse group of providers across services systems who often did not work together. However, each of the partners identified did provide services to or work with this triply burdened population. Thus, each had an incentive to collaborate on these efforts. The group included members BMC Division of Psychiatry, the Boston Public Health Commission, Vinfen Corporation, the Boston University School of Public Health, Consumer Quality Initiatives (CQI), the BMC Section of General Internal Medicine Primary Care Clinic, and Massachusetts DMH. Vinfen is one of the largest providers working with the DMH in Massachusetts. They currently had several programs which were designed to meet the needs of this population but recognized that this project provided a creative opportunity to truly address many of the challenges they faced in their existing programs.

BMC Division of Psychiatry and the DMH were well aware that the success of this program depended in large part on the ability of the program to meet the self-defined needs of the people the program was designed to serve. While clinical and policy experience, as well as the research literature, each provided some insights into these needs, a decision was made to partner with the consumer community, early on, in the development of the proposal, to best assure that the project was consumer-informed. Consumer involvement has been a major component in the development and evaluation of the BMC ACCESS Project. The idea of expanding Safe Haven options in Boston came from the relationship between DMH and homeless advocates. The DMH-funded organization Homeless Empowerment Advisory Project, comprising former and current homeless individuals, was involved in the development and implementation of the BMC ACCESS Project. In addition, the group partnered with CQI, a consumer-led, nonprofit organization whose board of directors is made up of 51% mental health consumers. CQI provided consumer input on all aspects of the

development and evaluation of the Dudley Inn. CQI members serve on all committees and were involved in the design and conduct of all evaluation activities.

BMC ACCESS Project: Program Components

The enhanced model developed offers mental health services, substance abuse treatment, and referral and primary care on-site in the Safe Haven. This model was developed to reduce barriers to these services, provide seamless care, and address the high level of need for a diverse array of services among the residents. Importantly, in keeping with the "low barriers to entry" approach used in the Safe Haven model, residents are not required to participate in services; however, services are consistently offered and available.

Housing

Three critical characteristics of the housing site chosen were identified. First, its central location and proximity to BMC and public transportation was key in outreach efforts to potential residents and in working with residents to maintain their existing support services. Second, in keeping with the Housing and Urban Development Safe Haven model,[9] residents of the Dudley Inn have their own rooms and keys. This is critical for outreach but also for maintaining a sense of individual safety and respect within the house. Given the high level of paranoid symptoms among the residents of this program, security and safety issues were addressed early on in the physical design of the space. Finally, securing a site with multiple floors, each with its own common area, allowed us to serve men and women on separate floors thereby creating safe spaces for all.

Outreach

The initial partnership between the BMC Division of Psychiatry and the Metro Boston Region of the Massachusetts DMH allowed for the project to have access to a wide range of clinical and other support services. The outreach for the Safe Haven is conducted by the DMH Homeless Outreach Team (HOT). As part of their existing mission, HOT identifies people who are chronically homeless and have historically refused all services, including generic and DMH shelters and other outreach efforts. While the HOT had identified these people in need of services, there were few actual services appropriate for this population, and so the creation of the Safe Haven filled an important gap for the workers. All potential residents are assessed for eligibility by the HOT. Eligibility criteria included: chronic homelessness, street dwelling, active substance use, and a severe mental illness. In addition, "fit" for living in the program was assessed by evaluating potential for danger to themselves, house staff, and other residents.

Clinical model of care

The clinical model of care is designed to provide the elements of care consistent with the needs of this particularly vulnerable population. Essentially, this population's care needs require a number of service elements including psychiatric, medical, housing, and substance abuse services. In addition to these individual elements, there is an obvious need for thorough coordination and efficient access to needed services and follow-up care. Vinfen Corporation is responsible for the 24 h a day/7 days a week staffing and daily operations of the program; they provide a director for the house as well as seven full-time staff (including both Master's level and Bachelor's level mental health counselors).

The care of the residents of the Dudley Inn is coordinated by a multi-disciplinary, multi-agency team representing the service elements described above. The team meets weekly at the program to

discuss cases and to provide services if the residents request it. The weekly meetings, with services on request, maintains a low-demand environment for the residents, and the team's presence at the program facilitates maximum support for the residential staff. It also provides an opportunity for the team to assess the residential milieu. The team includes: the residential Program Director; a psychiatrist; a substance-abuse case manager; a primary-care physician; and the Director of Program Management who oversees housing and homeless services, for the Metro-Boston Area Office of DMH. The integration of care if further supported by the use of one medical record/chart for each resident which includes all services received on-site as opposed to the numerous and different paperwork and charting required by each of the partners. Integration and coordination are also enhanced through informal mechanisms such as through electronic and telephone communication regarding relevant issues that develop within the program.

Psychiatric clinical activities include continued assessment of Dudley Inn residents in a way that emphasizes observation rather than intrusive questioning. Similarly, treatment is offered, and efforts are made to minimize even the suggestion of coerciveness. As such, treatment such as pharmacological management is provided as requested by the residents. Consistent with rehabilitation and recovery practices, residents are invited to actively engage in their treatment planning. This includes asking their input regarding whether their experiences with a particular medication regimen has been problematic in the past and/or carefully exploring with the residents the expected benefits and possible side effects from the medications. The residents' involvement with treatment planning also includes assisting staff in identifying strategies for maintaining safety within the program. In anticipation of a resident not being able to safely maintain themselves and others, all staff are trained in de-escalation techniques, and every effort is made to maintain this model. However, if an imminent risk to either the resident or others is evident, treatment may be required, or involuntary hospitalization may be used. Psychiatric support also includes provision of 24 h/day, 7 days/week clinical coverage to the program. This has allowed for greater confidence in the low-demand model by all staff.

The *substance abuse* case manager provides ongoing assessment of each resident's readiness for change and provides valuable linkages to the range of substance abuse services available. By using evidence-based, motivational interviewing techniques, the clinician is able to approach each resident's ambivalence in a way that is consistent with the low-demand model as opposed to an explicit or implied expectation of abstinence. Thus, motivational interviewing as a clinical technique is superimposed on a harm-reduction theoretical perspective. In focusing on the harm associated with drug use without requiring a reduction in consumption, the resident is encouraged to focus more on the spectrum of harms than on abstinence. The substance abuse case manager and the entire clinical team also focus efforts at harm reduction to those behaviors that may compromise housing viability.

Medical activities are similar in approach and scope to the psychiatric clinical activities. The primary-care physician observes for evidence of significant medical illnesses, particularly those that are potential complications of long-standing substance use. The physician communicates and coordinates care with the larger medical community. The physician builds trust with residents gradually over time through engagement over minor medical concerns (e.g., urinary tract infections, upper respiratory tract infections, earwax removal, clipping toenails) and referrals for various conditions (e.g., dental, dermatological, rheumatologic, etc.). This low-intensity care enhances the physician's later ability to serve as the bridge between often highly independent, psychiatrically impaired residents and a medical system with limited capacity to adapt to these patients' specific needs.

The DMH Director of Program Management assists residents in applying for *housing*. In addition, the Director of Program Management serves as a critical link to the full array of services within the DMH system. The Homeless Outreach Team Director remains involved in the clinical team after residents transfer to the Dudley Inn as well as after they leave the Dudley Inn providing

consistency for residents as well as continuity to the team and help in formulating what type of permanent residential setting may be best for the resident.

Finally, if residents have relationships with community providers, they are invited to join the Safe Haven clinical team for meetings with the resident. Again, this helps to ensure continuity and capitalizes on existing trusting relationships between the resident and service providers. The use of this integrated services model provides an intensive-care approach delivered in a residential setting.

Transition to permanent housing

Discussions of permanent housing begin very early in the resident's stay and starts with the residents' hopes and dreams for housing. Staff members discuss the types of housing residents would like and the neighborhoods they would prefer to be in. Then the work begins to pull together the resources available to the BMC ACCESS team to try to develop the best match with the residents. The discharge planning process maintains the focus on behaviors that potentially compromise housing as plans are developed. The first step in the discharge process is gaining a thorough understanding of the behaviors that may or may not be related to their psychiatric illness and/or substance abuse that will likely interfere with both acquiring and maintaining a permanent residential setting. The team then works with residents on those behaviors and how to manage those behaviors.

Metro Boston DMH has a waitlist to receive housing placement; however, Dudley Inn residents have priority placement status and have the full range of housing options available to them. These options range from those with staff supervision to those that are independent with support from a case manager. Finally, the discharge planning includes a process that incorporates the community personnel in the process well before the resident is actually discharged from the program. This includes both those with whom the resident has existing relationships and any staff affiliated with programs to which the resident is moving. This allows the resident to know the person that they will be engaged with in the community and for the community personnel to fully understand the perspective on the resident that has been developed during the stay at the Dudley Inn. An eviction-prevention planning framework is used. Discharge planning that is focused on housing rather than treatment has proven to be the best approach for meeting the challenge of acquiring and maintaining housing for this challenging group of residents for whom a treatment-centered approach has not been effective.

Evaluation of the BMC ACCESS Project

The consumer-informed evaluation was a critical component of the BMC ACCESS Project. The data generated through the evaluation process was used by the leadership and clinical staff to consistently examine and adapt the model of services being provided. In addition, the evaluation data were actively used as part of the sustainability efforts of the grant-funded project. In fact, these data proved critical in the success of these efforts. The project has now been sustained as the Safe Haven has been incorporated into the Massachusetts Department of Mental Health's array of services and is fully funded through the Department of Mental Health. The evaluation component of this project takes on increased importance given the lack of evaluation data that exists related to the development of Safe Havens for the dually diagnosed population and particularly the impact of integrated mental health, substance abuse, and primary-care services. A major implementation challenge of conducting evaluations of these types of programs was addressed in the design. That is, the need to balance the evaluation design, with the low-demand philosophy of the Safe Haven treatment modality. The partnership with CQI and other consumer partners helped to facilitate this process. Consumer Quality Initiatives' evaluation responsibilities include conducting the

qualitative interviews and disseminating reports to key public policymakers, provider organization, and advocacy groups.

The evaluation design is longitudinal and relies on collecting both qualitative and quantitative data at five time points over a 1-year period. All residents of the Dudley Inn are eligible to participate in the evaluation. The participation rate is 94.7% at baseline (18 of 19). All residents are informed during outreach activities and again upon entry that the Dudley Inn has a research component and that their participation is entirely voluntary. This ensured that residents were not surprised by activities in the Safe Haven. After spending two consecutive nights in the house, the resident is approached by the staff about participating in the research study, and after verbal consent is given, research staff is contacted to get written consent and schedule interviews.

The primary outcomes examined are housing trajectory and level of engagement. Success was measured by two criteria: (1) do people stay housed? and (2) are they connected to services, social networks, and/or supports in the community? A logic model of the BMC ACCESS Project that describes the program activities and short-term, intermediate, and long-term outcomes is presented in Figure 1. A detailed description of the evaluation of the BMC ACCESS Project is forthcoming; however, here the demographic characteristics of the residents of the Safe Haven at baseline are presented. Table 1 describes the major characteristics of the residents who have lived in the Dudley Inn and participated in the evaluation (N=18). Only one resident chose not to participate. In

Figure 1

BMC ACCESS Project logic model

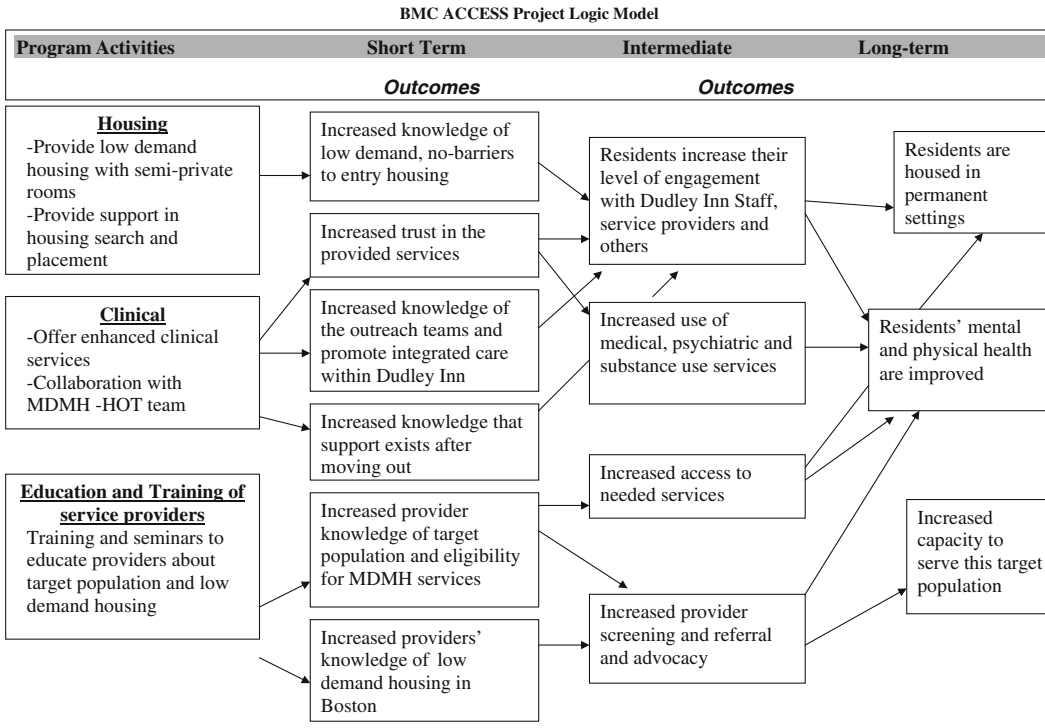


Table 1
Baseline characteristics of Dudley Inn residents (N=18)

Variable	Number	Percent
Gender		
Female	8	44.4
Male	9	50.0
Other	1	5.56
Age		
30–40	3	16.7
41–50	9	50
51–60	4	22.2
≥61	2	11.1
Race		
Black or African American	9	50
White	6	33.3
Native American	1	5.6
Other	2	11.1
Highest education completed		
8th grade or less	2	11.1
Some high school	7	38.9
12th grade/high school	8	44.4
College/university	1	5.6
Years of homelessness prior to living in the Dudley Inn[a]		
2–5	5	38.5
6–10	3	23.1
11–20	4	30.8
21+	1	7.7
Trauma history[b]		
Any experience of assault in the past	9	56.3
Sexually/physically assaulted by someone close	7	43.8

[a]Based on N=13
[b]Based on N=16

addition, the discussion of "lessons learned" is supported with exemplary quotations from the qualitative interviews with residents of the Safe Haven.

Lessons Learned from the BMC ACCESS Project

Many lessons have been learned by the people involved in this project throughout its course. They are described here as they relate to the program development, program components, and the evaluation. The most important lesson has been the reminder that many of the struggles faced by clients do not derive from internal limitations but are either exacerbated or driven by limitations of the system of care. Many policies and procedures required to access services present unassailable barriers to people with long histories of homelessness, severe mental illness, and histories of substance use. Only by systematically identifying these barriers and removing them were people able to begin to transform their lives. Barriers removed included not requiring sobriety or treatment

adherence for housing and, in addition, not requiring Medicaid Rehab Option-compliant treatment plans, training protocols required by some agencies, and other funding-specific process-adherence measures and performance-based contracting requirements because of the unique nature of the program.

Development of the BMC ACCESS Project

Several lessons should be emphasized from the development of the BMC ACCESS Project. Much of the success of the project was due to strong interpersonal relationships across institutions. Each of the partners took on risk with developing this new model of services, and this was facilitated by a history of trusted working relationships.

In planning the project, the number of people anticipated to be served annually was overestimated. Due to the lack of evaluation data on Safe Havens, the initial proposal was to serve about 15 people per year with the eight-bed Safe Haven. These numbers were based on a belief that about 1/3 of the people who entered the Dudley Inn would "blow themselves out." The team of experienced clinicians and researchers believed that many of the people coming in would not be able to tolerate the milieu. In fact, this has become a challenge in working with SAMHSA, as all but two of the people who have entered the Dudley Inn have stayed until they moved to either a hospital or permanent housing. Working closely with the HOT team has enabled residents to maintain trusting relationships during the transition in, with street outreach workers they have known for years. In addition, the policy only requiring people to sleep in for two nights a week during the transition allows people to enter in a way that is comfortable for them. This success in helping people to remain indoors meant that the target number of people served during the grant tenure was not met.

A multi-system collaboration presents challenges in developing project policies and procedures since each collaborating partner had their own values, policies, and interests. This was evident in a broad spectrum of planning and implementation discussions, including issues such as what type of record keeping should be used for residents of the Safe Haven, the development of house rules, how to maintain Health Insurance Portability and Accountability Act compliance with so many separate entities participating in providing care, and multiple Institutional Review Boards (IRBs) for the research activities.

Finally, subcontracting with a traditional residential services provider has highlighted several of the ways in which the Safe Haven model is vastly different from residential services for people with severe mental illness. Establishing nontraditional program policies required the flexibility of each of the partner agencies and the shared commitment to providing a truly low-demand, person-centered, and high-quality service. Oftentimes, this required the provider to turn existing rules on their head. For example, Department of Mental Health regulations require Vinfen programs to take a picture of all residents upon entry into the program. This photograph is then placed in the client record so that it is available if needed. Residents entering the Dudley Inn were often highly paranoid and struggling with the transition in, as described below. In the Safe Haven, residents were not required to have their picture taken, and instead pictures were required to be taken of all project staff who spend time in the house. These pictures are posted on a bulletin board outside the front office, with descriptions of who each person is and their role on the project.

Program components

The design and layout of the brownstone and the neighborhood were not initially considered to be program components, but in fact, they became the first two critical program components. The neighborhood was attractive for some residents and problematic for a few residents who were not comfortable at first in a diverse city neighborhood. The Safe Haven is located on a residential street

lined with brownstones, off of a busy urban square with many small businesses and access to public transportation. As one resident stated:

> We're right in the center of transportation, and you can walk right down the Avenue. The package store's right there, and you can really walk into the station. There's a couple package stores, there's a pizza shop, and clothes, and brand new small business. And it's very convenient, and it's very very nice. The people have been fairly decent. Their behavior in this neighborhood has been fairly decent.

The Safe Haven is in a brownstone building with four levels. Each floor has several bedrooms and a common living space. The lower level has no bedrooms but has a kitchen and a living room for residents. Traditional residential programs have a locked staff office often in the back of the program. In designing the Safe Haven, the staff office was moved to be directly off the front entrance, and whenever possible, the door is open. There is no front desk where residents have to check in; however, residents do need to ring a buzzer to be let into the building for security reasons. This allows residents to walk into the Safe Haven and go directly up to their rooms if they would like. Dudley feels more like a rooming house to residents than a shelter and, therefore, does not feel institutional. A resident said, "I think the fact that it is an apartment building. And it's not like an institution—[not] like 12 bedrooms on each side and a long hall and the nursing station. I think that's what made me feel more comfortable, more homey."

A key feature of the Safe Haven is that each resident gets his/her own bedroom which they can lock with a key. Having a private bedroom was a huge benefit for many of the residents. One person said "having that comfortable bed" was the most helpful thing during the transition in. Another said that what made him most comfortable in the house was "The bed. Going to sleep. Not having to listen to people snore and cough all night, talking, grinding their teeth. It's a mad house—(another shelter) is a madhouse. You got people crawling all over you." Having a locked room also gave people a sense of security and safety that was rare for them. One said, "I just don't feel insecure. My property is protected by one little lock on the door, and we all have the same benefit here. All eight of us have our own private area where we can keep our little treasures and no one can violate that." Another resident said, "You got your own door keys. You can come and go as you please." Having a safe, secure place of their own was very valuable to residents.

The staffing plan changed over the course of the project. The initial proposal included several spots designated for master's level clinicians as well as the premise that there would be a master's level clinician in the house at all times. The unique model of this project created staffing challenges. Often, the sought-after highly educated staff had been trained in other models or philosophies of providing care or had worked in programs with other frameworks. It was quickly learned that the qualities of successful staff extended beyond clinical expertise and experience and included a high level of tolerance for unusual behaviors, flexibility, creativity, and in a core way, people who liked the clients. It is critical that staff continually re-examine the model that services are voluntary for the residents, but it is required that staff consistently offer them. While ideally master's level people with these characteristics were found, these strengths were often found in staff with less formal education.

Aside from the multidisciplinary nature of the team, there are four characteristics of the professional and paraprofessionals that are crucial for this model of care. Caregivers must be *supportive* of the residents' efforts including those that in some other settings might be viewed as minor. Given that many of the residents have not engaged with any caregivers in any way, the mere fact of their willingness to stay in the program is worthy of support. Staff *adaptability* is a desired characteristic. Caregivers with a desire to go beyond their formal job descriptions have been highly successful in this setting. For example, the psychiatrist might help the resident do dishes as opposed to doing a formal interview. Because all members of the team have their own clinical and theoretical framework and because the residents' range of difficulties might not fit neatly into those frameworks, staff must be open to examining their frameworks and integrating their experiences

with the program. *Resourcefulness and tolerance* are the final characteristics that have found to be consistent in all members of this clinical team. Team members have been resourceful enough to identify previously unknown resources for residents or identify means of accessing medical care for residents who cannot comply with the usual referral process for a medical system.

Many residents of the Dudley Inn are stably housed for the first time in decades. There were many difficulties and obstacles identified in working with residents to transition on to permanent housing options. Some residents would prefer to stay at the Safe Haven, having worked hard through a difficult transition of coming in and having built trusting relationships with staff and other residents. For these residents, maintaining a connection with the staff and former residents while they are in the community has been very helpful. In addition, the transition from the Safe Haven to other housing is a very stressful one, and it requires that the staff work closely with community-based providers to assure a high level of continuous care. Maintaining the relationships between residents and the HOT workers who assisted in outreach to the program has been critical as residents transition to next-step housing. Finally, eviction-prevention plans are developed with each resident in order to increase the likelihood that each person will remain in their next-step housing solution.

Evaluation

Several important lessons can be drawn from experience to inform future efforts to evaluate Safe Havens or other transitional housing models for this population. First, having consumer and other stakeholder involvement in every aspect of the evaluation was challenging but certainly a critical investment. The evaluation team included several consumer members as well as members from each of the partner organizations and researchers. In addition, the use of trained consumer interviewers has facilitated building a rapport with participants and contributed to the ability to follow participants over time. High follow-up rates for the quantitative interviews were obtained because of the relationships built between the evaluation team and the residents. Residents reported feeling respected by the research interviewers and enjoyed having a chance to tell their stories.

The principles of community-based participatory research (CBPR) were relied upon in this project, and this created several challenges. First, as many have noted, the pace of CBPR is slower than traditional research and evaluation. First, the team educated themselves about each other's experiences and areas of expertise. This cross-training, while time consuming, allowed us to use a shared language and to acknowledge the breadth of knowledge shared by the group. As many evaluation decisions required a consensus-building process, the reliance on these principles often made it difficult to meet deadlines set by the funder. Another challenge was scheduling evaluation meetings. The wide range of members, with varying schedules, made this particularly difficult. The diverse evaluation team has been beneficial in several key ways including: the design of sensitive and respectful research instruments, a thoughtful consent process, the gathering of rich data by trained consumer interviewers, targeting dissemination activities to allow for efficient adoption of change, and in sustainability efforts.

Secondly, the reliance on a mixed-methods design has proven to be a very useful strategy. Few structured instruments have been validated in this population, and while several instruments appear to have face validity, more work is needed to understand the meaning of these instruments for the population served by this project. As the numbers of participants is still relatively low, the qualitative data were used to revise the structured instruments to focus on appropriate areas, to help us interpret trends in the quantitative data, and to provide us with a rich view of the perspectives and experiences of residents which would not be accessible through standard structured instruments.

Third, working as an inter-agency collaboration necessitates abiding by the policies and procedures of each partner. This creates many challenges, as often these can be conflicting. This was certainly the case in dealing with the multiple IRB approvals required for this project, as often

what one agency required for language on the consent form was not acceptable to another agency. In order for truly collaborative work to be supported, efforts must be made by each agency to accommodate the needs of the other partners.

Finally, working with people who have long histories of chronic homelessness, substance abuse issues, and severe mental illness challenges us to struggle with the concept of coercion in research. This was addressed in several ways. The outreach team informed potential residents that research was going on in the Safe Haven as they conducted outreach activities on the street. Introducing this as early on in the process as possible both allowed participants time to think about the research components of the project but also prevented people from being surprised upon arrival by questions from researchers, especially given the low-demand philosophy of the program. In addition, the evaluation team carefully considered the use of research incentives and learned that it was important to inform the Dudley Inn staff when these incentives were provided to participants. Participants were also informed that this would be done. Perhaps the most important efforts were made around making a clear distinction between the provision of services (housing) and participation in the evaluation. At every interview, the consent form was reviewed, and residents were reminded that participation was voluntary and that they could choose to stop participating at any time and that it would not impact anything about the stay at the Safe Haven. Identifying staff who provided services and research staff in photographs on the bulletin board and making every effort to be clear when activities were part of the evaluation.

Conclusion

In this paper, the development of a multi-agency, consumer involved, medically enhanced Safe Haven is described. Residents of the Safe Haven bring a vast history and set of experiences with them as they transition into services. For programs to effectively partner with people as they seek safety or to transform aspects of their lives, they must be flexible in the ways they provide care.

A model which provides the highest level of care, through integrated services, experienced clinicians, leadership from each of the collaborating agencies, and person-centered, recovery-focused care, to those most in need challenges the norms and assumptions in the ways systems of care are constructed. However, when services are provided in this manner, people are able to transform their lives in unexpected ways that have led to them to end their time on the streets and transition to permanent housing solutions.

Implications for Behavioral Health

There are several critical implications of the experience of the BMC ACCESS Project for mental health and housing practice and policy. The first is the need to identify and reduce regulatory and financial barriers to developing innovative programs. These include policies and regulations which categorize programs as either “residential facilities” or “shelters.” Each of these categories brings with them licensure requirements, and the unique nature of Safe Havens makes it difficult to appropriately categorize the program. In addition, operating the Safe Haven is quite similar in cost to operating a traditional DMH residential facility for a similar population. However, this Safe Haven is more costly due to the enhanced clinical services which are provided on-site. As efforts to sustain the clinical activities are made, the attention has focused on the need to pursue the development of models of reimbursement for mental health and primary care, which focus on the person served, as opposed to distinct clinical activities. It is clear that a high level of care is provided in the Safe Haven; however, there are few billable clinical interactions to sustain the activities of the enhanced clinical team.

Work is needed to develop policies and procedures, both locally and at a higher level, which allow for a reinterpretation of how to assess people struggling with mental illness and develop

treatment plans. In this project, traditional models were not relied upon, and it is important to explore how the lessons learned with the Dudley Inn might best serve the needs of people seeking care in other programs and treatment modalities within the Division of Psychiatry and other services. One clear need is to examine where housing is placed as a priority in the approach to people with severe and persistent mental illness.

Despite two decades of subsidy programs designed to house the homeless, there are not enough leased housing dollars in the system to meet the needs of all homeless people. Consequently, the most disenfranchised, e.g., people who are both mentally ill and addicted to substances with protracted histories of "on the street" homelessness, have not been reached by the current system of support. The Housing First approach has been shown to be effective in reducing homelessness for people with serious mental illness and/or substance addiction; however, the success of the Medically Enhanced Safe Haven model points to the efficacy of a transitional step for this specific subset of the homeless population. There are several reasons why a transitional step for this population is often advantageous. First, there are a number of practical barriers to accessing permanent subsidized housing for this subset of the homeless population, including the need for an address, personal identification, and recent proof of income in order to apply for subsidized housing which most people entering the Safe Haven do not possess. Further, people who have lived on the streets usually have open legal charges which must be resolved in order for housing agents to accept their application. Also, most people in this subset of the homeless population have refused all state and federal entitlement support and therefore have no way to pay for food or utilities if they are housed. The Safe Haven allows for these practical barriers to be worked on before application is made to permanent housing, greatly increasing the likelihood of a successful application.

Second, the personal characteristics of this subset of the homeless population pose challenges for them in obtaining and keeping housing. These include a general suspicion of "helpers" and lack of trust that "authorities" will work to meet their needs on their terms. The environment of the Safe Haven is therapeutic in fostering trust with helpers first with staff and then within the larger system of care. Further, this population has a propensity to be fiercely independent—therefore they abhor the simple act of signing a lease which obligates them to act in ways acceptable to others and may leave housing if/when issues arise instead of cooperating with others toward a mutually acceptable solution. The work of the Safe Haven is to help the residents to see that it's worth bothering with/believing in/buying in to needed supports. This population often engages in a range of behaviors inconsistent with housing stability, e.g., intrusiveness with other tenants, drunk and disorderly behavior, screaming/responding to hallucinations late at night, urinating in public, acting on paranoia re: other tenants with intimidating and/or aggressive behavior, etc. The environmental supports of the Safe Haven allow them to process incidents and practice more adaptive behaviors which increase the likelihood of their success in permanent housing.

Lastly, the extent and nature of the medical conditions presented by this subset of the homeless population make a "direct to permanent housing" less likely to succeed. People served by the Safe Haven over the last 5 years have experienced renal failure, breast cancer, advanced cirrhosis of the liver, stroke, and advanced cellulites in addition to the usual medical consequences of mental illness and substance addiction. Since the Safe Haven has on-site medical and psychiatric physicians who provide both primary care and access to acute services with their continued coordination and support, residents have addressed long-standing medical issues for the first time, thereby increasing the likelihood for successful community placement instead of nursing home placement.

Finally, difficulties with helping people transition to "more permanent" housing options raise the question as to whether there is a need to consider expanding the model of a continuum of housing options to include more permanent Safe Haven beds or other types of programs which allow people who have been living on the streets for long periods of time to be engaged, come in, and then stay

where they feel safe and at home, without the need to move to a permanent housing option. These experiences and evaluation data cannot answer this question but certainly highlight the need for some discussion in this area.

Acknowledgment

The Boston Medical Center (BMC) ACCESS (Advanced Clinical Capacity for Engagement, Safety and Services) Project was supported by a grant from the Substance Abuse and Mental Health Services Agency (TI14028).

References

1. Anderson D, Fernandes K, Green G, et al. *Homelessness in the City of Boston Winter 2002–2003, Annual Census Report.* Available at: http://www.ci.boston.ma.us/shelter/pdfs/report_02.pdf. Accessed April 6, 2006.
2. National Coalition for the Homeless. *NCH Fact Sheet #5: Mental Illness and Homelessness*. June 2006. Available at: www.nationalhomeless.org. Accessed March 31, 2007.
3. Shern D, Tsemberis S, Anthony W, et al. Serving street-dwelling individuals with psychiatric disabilities: outcomes of a psychiatric rehabilitation clinical trial. *American Journal of Public Health*. 2000;90(12):1873–1878.
4. Dennis DL, Buckner JC, Lipton FR, et al. A decade of research and services for homeless mentally ill persons. Where do we stand? *American Psychologist*. 1991;46(11):1129–1138. doi:10.1037/0003-066X.46.11.1129.
5. Fischer PJ, Breakey WR. The epidemiology of alcohol, drug, and mental disorders among homeless persons. *American Psychologist*. 1991;46(11):1115–1128. doi:10.1037/0003-066X.46.11.1115.
6. Hwang S. Mental illness and mortality among homeless people (Editorial). *Acta Psychiatrica Scandinavica*. 2001;103(2):81–82. doi:10.1034/j.1600-0447.2001.103002081.x.
7. Rosenheck R. Cost-effectiveness of services for mentally ill homeless people: the application of research to policy and practice. *American Journal of Psychiatry*. 2000;157(10):1563–1570. doi:10.1176/appi.ajp.157.10.1563.
8. Ward Family Foundations. *Safe Havens: Analysis of Strategy and Operating Practices.* July 2005. Available at: http://www.wardfamilyfoundation.org/shp.shtml. Accessed August 31, 2005.
9. U.S. Department of Housing and Urban Development. (2001). *In from the Cold-Safe Havens for Homeless People*. Accessed at http://www.hud.gov/offices/cpd/homeless/library/havens/index.cfm. Retrieved April 12, 2006.

J Thromb Thrombolysis (2008) 26:167–174
DOI 10.1007/s11239-008-0272-4

Antithrombotic therapy in atrial fibrillation: guidelines translated for the clinician

Renato D. Lopes · Jonathan P. Piccini · Elaine M. Hylek · Christopher B. Granger · John H. Alexander

Published online: 21 September 2008

Abstract Atrial fibrillation (AF) is the most common cardiac arrhythmia in clinical practice, accounting for approximately one-third of hospitalizations for cardiac rhythm disturbances. The highest incidence of AF is in patients 70–80 years old and other high-risk populations. Although the diagnosis of AF is usually straightforward, effective treatment strategies are less well implemented. This is particularly true for antithrombotic therapy, which is very effective at preventing thromboembolic complications of AF. Stroke is the most significant morbidity in AF patients. The yearly risk of stroke increases from 1.5% for AF patients aged 50–59 to 23% for those aged 80–89. Ischemic strokes secondary to AF carry twice the risk of death when compared with strokes from other causes. We provide a practical and useful review of the most recent American College of Cardiology/American Heart Association/European Society of Cardiology guidelines-based care and future directions of antithrombotic therapy for patients with AF.

Keywords Atrial fibrillation · Antithrombotic therapy · Guidelines

Fred Spencer, MD, Professor of Medicine, McMaster University served as the Guest Editor for the manuscript.

R. D. Lopes · J. P. Piccini · C. B. Granger · J. H. Alexander (✉)
Duke Clinical Research Institute, Duke University Medical Center, Box 3850, Durham, NC 27710, USA
e-mail: John.H.Alexander@duke.edu

E. M. Hylek
Department of Medicine, Boston University Medical Center, Boston, MA, USA

Atrial fibrillation

Atrial fibrillation (AF) is a supraventricular tachyarrhythmia characterized by uncoordinated atrial activation with consequent deterioration of mechanical function [1]. AF is the most common cardiac arrhythmia in clinical practice, accounting for approximately one-third of hospitalizations for cardiac rhythm disturbances. The estimated prevalence of AF is 0.4 to 1% in the general population, increasing with age, such that nearly 1 in 10 people older than 80 years have AF [2, 3]. Accordingly, the highest incidence of AF is in patients between 70 and 80 years of age (Fig. 1) [4]. The incidence of AF is even greater in high-risk populations, such as patients with heart failure (HF) [3].

The burden of AF in the United States is increasing; by the year 2050 there will be an estimated 12.1 million (95% confidence interval [CI] 11.4–12.9) (2.4-fold increase from 2000) Americans with AF. However, this estimate assumes no further increase in the age-adjusted incidence of AF beyond 2000. If the incidence of AF increases at the same pace, then the projected number of adults with AF would be 15.9 million, a 3-fold increase from 2000 [5].

As demonstrated in Fig. 2 [5], AF is a pervasive and growing public health problem with significant socioeconomic implications. AF is associated with increased risk of stroke, HF, cognitive dysfunction, and a 2-fold increase in mortality [5, 6]. The rate of ischemic stroke among patients with nonvalvular AF averages 5% per year, 2–7 times that of the general population [7]. The risk of stroke increases from 1.5% for patients with AF 50–59 years old to 23% for those 80–89 years old [8]. Systemic hypertension, diabetes mellitus, HF, myocardial infarction (MI), and obesity are considered risk factors for AF as well as independent risk factors for stroke. Aggressive primary prevention and intervention once these risk factors are present are essential

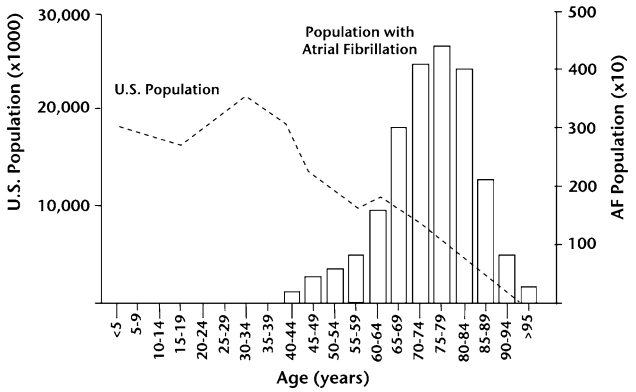


Fig. 1 Age distribution of people with AF compared with the general population. (Reprinted with permission from Feinberg WM et al. Arch Intern Med 1995;155:469–73)

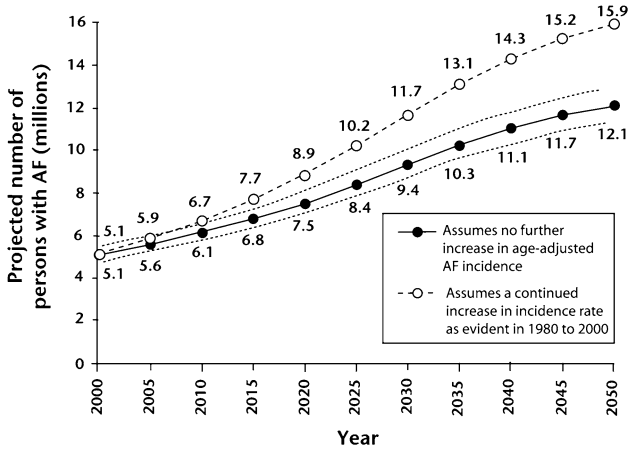


Fig. 2 Projected number of persons with AF in the United States between 2000 and 2050. (Reprinted with permission from Miyasaka Y et al. Circulation 2006;114:119–25)

if we are going to successfully contain this epidemic. In addition to curbing the incidence of AF, attention must also be directed to preventing morbidity and mortality in patients who already have developed AF. Although the diagnosis of AF is usually straightforward, treatment strategies are less well defined, particularly antithrombotic therapy, which has an enormous role in preventing thromboembolic events.

The American College of Cardiology/American Heart Association/European Society of Cardiology guidelines for AF [1] recommend antithrombotic therapy based on the number of risk factors. Risk factors associated with a high risk of stroke (>5% per year) include previous stroke, TIA, systemic embolism, mitral stenosis, and the presence of a prosthetic heart valve. Moderate risk factors (stroke rate 3–5% per year) are age ≥75 years, hypertension, HF, and diabetes. Less validated or weaker factors include female sex, age 65–74 years, coronary artery disease, and thyrotoxicosis.

Antithrombotic therapy (which includes both antiplatelet and anticoagulant agents) should be given for all patients with AF, except those with lone AF (<60 years without heart disease or risk factors for stroke) or contraindications (Class Ia). The selection of the antithrombotic therapy should be based on clinical judgment, taking into account risk of stroke and bleeding (Class Ia). Monitoring of the international normalized ratio (INR) should be done at least weekly during the initiation of anticoagulation therapy and monthly when the anticoagulation is deemed stable (Class Ia). Although a different arrhythmia from AF, patients with atrial flutter have the same indications for antithrombotic therapy as patients with AF (Class Ia).

Patients with AF and no risk factors for stroke should receive aspirin (81–325 mg/day) (Class Ia). Patients with 1 moderate risk factor (age ≥75, hypertension, HF, diabetes mellitus, left ventricular ejection fraction [LVEF] ≤35%) or any low risk factors should receive aspirin (81–325 mg/day) or warfarin (INR 2.0–3.0) (Class IIa) depending on patient preference, bleeding risk, and ease of monitoring the INR. The recent American College of Chest Physicians (ACCP) guidelines concur with this recommendation, but with a suggestion that these intermediate risk patients should preferentially receive warfarin rather than aspirin [9]. Patients with any high risk factors (prior stroke, TIA or embolism, mitral stenosis, prosthetic heart valve) or more than 1 moderate risk factor should receive warfarin (INR 2.0–3.0) (Class Ia).

The Major risk factors for ischemic stroke and systemic embolism in patients with nonvalvular AF are previous stroke or transient ischemic attacks (TIA), diabetes mellitus, hypertension, HF, and advanced age. These risk factors are the elements that form the $CHADS_2$ score [10, 11]. This score, which is a simplified algorithm of major and moderate risk factors for stroke, ranges from 0 to 6, is easy to use in clinical practice, and provides a reasonable estimate of the risk of stroke in patients with AF (Fig. 3) [1]. The adjusted rates of stroke vary from 1.9% in patients with a $CHADS_2$ score of 0–18.2% with a $CHADS_2$ score of 6. The description of risk factors and recommendation of antithrombotic therapy according to risk factors and $CHADS_2$ score are summarized in Table 1.

The selection of antithrombotic therapy should be done using the same criteria regardless of the type (paroxysmal, persistent, or permanent) of AF (Class IIa). All patients with AF should be reevaluated regularly for the need for anticoagulation (Class IIa).

In older patients (≥75 years) with an increased risk of bleeding but without contraindication for anticoagulation, a lower INR target (range 1.6–2.5) may be considered for primary prevention of thrombotic events (Class IIb). Patients need to be informed of the trade-offs inherent to this strategy as studies have shown a greater frequency of

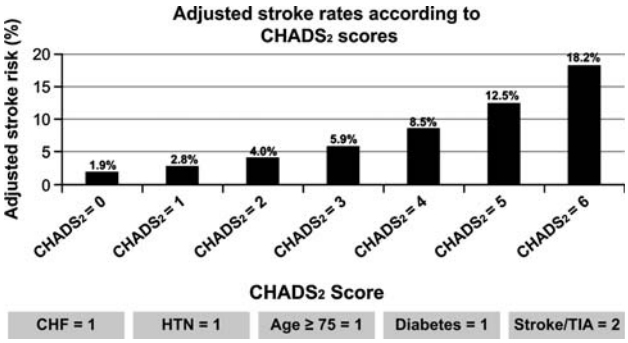


Fig. 3 Stroke risk in patients with AF not treated with anticoagulation according to the $CHADS_2$ score. (Adapted from Fuster V et al. Circulation 2006;114:e257–354)

stroke and more severe strokes when the INR is <2.0. The impact of this target range on bleeding is yet unclear [12].

Patients with lone AF do not benefit from aspirin, and the risk of bleeding in these patients is not well established (Class IIb). These patients should not receive anticoagulation therapy for primary stroke prevention (Class III).

Atrial fibrillation—bridging anticoagulants perioperative

There are few data to inform the question of thromboembolic risk during brief periods of warfarin interruption. In a recent study of 550 individuals with AF, 4 (0.7%) sustained a thromboembolic event in the 30-day period following an outpatient elective procedure [13]. When high-risk patients with AF undergo surgical procedures that will require interruption of anticoagulation therapy for more than 1 week, unfractionated heparin or low-molecular weight heparin may be used (Class IIb).

Atrial fibrillation postoperative

It is reasonable to use antithrombotic therapy in patients who develop postoperative AF, as recommended for non-surgical patients (Class IIa). Initiation of anticoagulation should be undertaken as soon as deemed safe from a surgical perspective.

Atrial fibrillation and acute coronary syndrome

AF following an acute coronary syndrome (ACS) is associated with worse short- and long-term clinical outcomes and more frequent in-hospital complications [14]. Anticoagulation with unfractionated heparin is a Class I recommendation. ACS patients post-stent placement represent a particular challenge in clinical practice, as they require dual antiplatelet therapy with aspirin and clopidogrel in addition to warfarin if additional risk factors for

Table 1 Antithrombotic therapy for patients with AF according to risk factors (Adapted from Fuster V, et al. Circulation 2006;114:e257–354)

Risk category		Recommended therapy
No risk factor		Aspirin, 81–325 mg daily[a]
At least one low risk factor	Female sex	Aspirin, 81–325 mg daily, or Warfarin (INR 2.0 – 3.0, target 2.5)
	Age 65 to 74 years	
	Coronary artery disease	
	Thyrotoxicosis	
Only one moderate risk	**C**ongestive heart failure[b]	Aspirin, 81–325 mg daily, or Warfarin (INR 2.0 – 3.0, target 2.5)
	Hypertension	
	Age ≥75 years	
	Diabetes mellitus	
At least one high risk factor	**S**troke or TIA	Warfarin (INR 2.0 – 3.0, target 2.5)
	Embolism	
	Mitral stenosis	
	Prosthetic heart value[c]	
More than one moderate risk factor		Warfarin (INR 2.0 – 3.0, target 2.5)
Lone AF[d]		No antithrombotic therapy recommended

$CHADS_2$ <1, Aspirin; $CHADS_2 = 1$, Aspirin or Warfarin; $CHADS_2$ ≥2, Warfarin. Doses as recommended above

[a] Patients with AF, but not lone AF

[b] Or left ventricular ejection fraction ≤35%

[c] If mechanical valve, target INR greater than 2.5

[d] Patients younger than 60 years without heart disease or risk factors for thromboembolism

stroke are present. Although the benefit of these medications for ACS and AF is well established, their use in combination is associated with an increased risk of hemorrhage [15–17] and not well defined. To date, there are few data to help guide physicians on how best to treat patients with significant risk for bleeding, such as patients with prior or recurrent gastrointestinal hemorrhage. A reasonable approach is to first stratify by patient stroke risk. If $CHADS_2$ is <2, then only aspirin and clopidogrel should be used. In patients with a $CHADS_2$ score ≥2, anticoagulation with warfarin (goal INR 2.0–2.5) in addition to low-dose aspirin (81 mg) and clopidogrel should be used. In this case, clopidogrel should be used for the shortest duration necessary, for example 4 weeks following bare metal stent placement. After clopidogrel is stopped, low-dose aspirin and warfarin (INR 2.0–3.0) should be continued long-term. The AF guidelines differ from the MI guidelines in respect to this point. The AF guidelines recommend that after clopidogrel is stopped, warfarin should be continued as mono-therapy. We believe, however, that based on the weight of the evidence regarding aspirin's benefit in coronary artery disease that warfarin and low-dose aspirin should be continued long-term. The duration of clopidogrel therapy will depend on the type of stent that is used. Bare metal stents require clopidogrel for at least 1 month while drug-eluting stents require at least 12 months of clopidogrel [18]. For this reason, use of bare metal stents should be strongly considered in patients with AF and ACS, particularly if they have a $CHADS_2$ score ≥2. The ACCP guidelines [19] also recommend frequent INR monitoring through an experienced anticoagulation clinic, if possible, and the use of a proton pump inhibitor, particularly among patients with risk factors or a prior history of gastritis and/or peptic ulcer disease. In addition, the ACCP guidelines recommend against "quadruple therapy" (heparin, warfarin, aspirin, and clopidogrel) unless the patient is at high risk of thrombosis and low risk of bleeding.

With respect to therapy of ischemic heart disease in patients with AF, the majority of patients can be safely managed with warfarin therapy alone [20]. Given the evidence supporting the benefits of aspirin however, for patients with AF and recent MI or percutaneous coronary intervention, combination therapy with warfarin and low-dose aspirin should be used, particularly if the patient is at low risk of bleeding. While both warfarin and aspirin are effective at preventing ACS-related ischemic events and AF-related stroke, warfarin therapy is more efficacious in the prevention of ischemic events than aspirin is in the prevention of stroke. Therefore, if the patient is at high risk of bleeding, aspirin should generally be stopped first, keeping the patient on warfarin alone with close follow-up. In all cases, the use of sound clinical judgment is paramount. Given the significant increased risk of bleeding [15–17], triple therapy (aspirin, clopidogrel, and warfarin) should be used for the shortest period possible. A corollary to this recommendation is that bare metal stents should be used instead of drug-eluting stents whenever possible in patients with AF who are prone to hemorrhage.

Atrial fibrillation and cardioversion

Patients with AF of 48 h or longer or with the duration of AF unknown should receive anticoagulation (INR 2.0–3.0) for at least 3 weeks prior to and 4 weeks after cardioversion (Class I). For patients who require more urgent cardioversion due to hemodynamic instability, an immediate anticoagulant effect can be achieved with unfractionated heparin initiated with a bolus injection followed by dose-adjusted continuous infusion to prolong the partial thromboplastin time to 1.5–2.0 times the reference control value. Oral anticoagulation with warfarin is recommended (INR 2.0–3.0) for at least 4 weeks after cardioversion (Class I). Because a significant number of individuals revert back to AF after cardioversion, indefinite use of warfarin or close surveillance of rhythm stability is essential. For patients with AF <48 h in duration and who are hemodynamically unstable, cardioversion should be performed without delay for prior initiation of anticoagulation (Class I). During the first 48 h after the onset of AF, the need for anticoagulation before and after cardioversion may be based on the patient risk of stroke (Class IIa).

The performance of transesophageal echocardiography (TEE) to assess the presence of thrombus in the left atrium or left atrial appendage is an alternative strategy that obviates the need for 3 weeks of warfarin prior to cardioversion (Class IIa). For patients with no identifiable thrombus, cardioversion can be done after anticoagulation with unfractionated heparin. Oral anticoagulation (INR 2.0–3.0) should be continued for at least 4 weeks after cardioversion (Class IIa). If a thrombus is identified by TEE and the patient is stable, warfarin should be given (INR 2.0–3.0) for 3 weeks prior to cardioversion and at least 4 weeks after restoration of sinus rhythm. A longer period of anticoagulation in this setting may be appropriate because of the high risk of thromboembolic events even after cardioversion (Class IIa).

All patients with atrial flutter undergoing cardioversion should follow the same indications for anticoagulation as for AF (Class IIa).

Atrial fibrillation and pregnancy

Antithrombotic therapy is recommended for prevention of thromboembolism throughout pregnancy for all patients

with AF (except for those with lone AF and low risk for stroke). The type of therapy should be chosen based on the stage of the pregnancy (Class I). Warfarin should be avoided in the first trimester due to teratogenicity and also during the final 4 weeks to decrease bleeding complications. Unfractionated heparin and low-molecular weight heparin do not cross the placenta and should be substituted for warfarin during the first trimester and final month or alternatively used throughout pregnancy (Class IIb). Warfarin use may be considered during the second trimester for patients with AF and high risk of stroke (Class IIb). Given the risks of bleeding in pregnancy and the lack of data in this population, antithrombotic therapy during pregnancy poses a particular challenge and should be approach on a case by case basis with input from practitioners with expertise in this area.

Atrial fibrillation and hypertrophic cardiomyopathy

Patients with hypertrophic cardiomyopathy and AF are at higher risk for stroke than the general population and should receive warfarin with a target INR of 2.0–3.0 (Class IIb), even with a $CHADS_2$ score of 1.

Future directions

While vitamin K antagonists, principally warfarin sodium, remain the standard of care for anticoagulation in AF, their use is limited by their narrow therapeutic window, marked interindividual variability in dose response, and numerous drug-drug interactions. Successful dose titration of vitamin K antagonists is essential in order to maximize efficacy and safety (Fig. 4) [1, 21, 22]; however, recent evidence suggests that patients spend nearly a third of their time outside their goal therapeutic window [23].

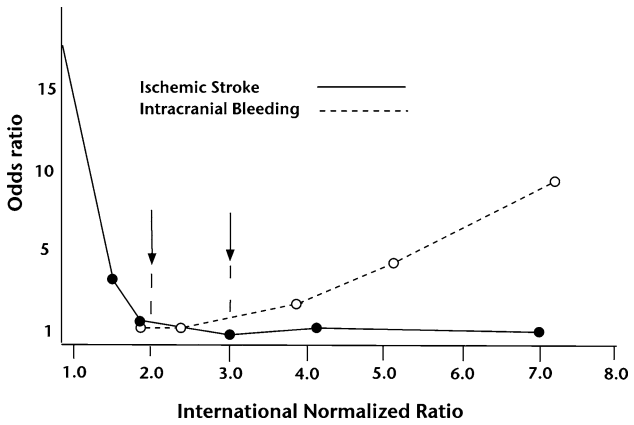


Fig. 4 Adjusted odds ratios for ischemic stroke and intracranial bleeding in relation to the intensity of anticoagulation. (Reprinted with permission from Fuster V et al. Circulation 2006;114:e257–354)

Accordingly, there is considerable need for and interest in the development of novel and more selective oral anticoagulants. Efforts directed toward the development of these agents have focused on direct thrombin inhibitors and factor Xa inhibitors (Table 2). The first direct thrombin inhibitor developed for use in stroke prophylaxis was ximelagatran. Unfortunately, ximelagatran while efficacious in the prevention of stroke, was associated with significant hepatotoxicity and was not FDA approved [24]. Dabigatran etexilate, the prodrug of the direct thrombin inhibitor dabigatran, is currently under development as an alternative oral anticoagulant. Dabigatran reversibly inhibits both free and clot-bound thrombin and has been shown to reduce deep venous thrombosis (DVT) in post-operative orthopedic patients when compared with enoxaparin (odds ratio [OR] 0.47, 95% CI 0.30–0.73, $P = 0.0007$), with a dose-dependent increase in major bleeding [25]. Dabigatran has a fast onset of action (Tmax 1.5 h) and is renally cleared [26]. At present, dabigatran is being evaluated for stroke prophylaxis in AF in a large phase III trial, the Randomized Evaluation of Long Term Anticoagulant Therapy With Dabigatran Etexilate (RELY) [27].

In addition to direct thrombin inhibition, there is also considerable work in progress to develop factor Xa inhibitors. Factor Xa is an attractive target as it is the common branch point for both the extrinsic and intrinsic coagulation pathways. Factor Xa inhibitors prevent thrombin formation by interfering with Xa upstream and, therefore, the prothrombinase complex [28]. The most recent factor Xa inhibitor evaluated for long-term anticoagulation in patients with non-valvular AF was idraparinux, a synthetic pentasaccharide whose long half-life permitted once weekly subcutaneous injection. The AMADEUS trial (Evaluating the Use of SR34006 Compared to Warfarin or Acenocoumarol in Patients With Atrial Fibrillation) was stopped early due to an excess of clinically relevant bleeding among patients randomized to idraparinux, including increased intracranial bleeding (1.1 vs 0.4 per 100 patient years, $P = 0.014$) [29]. Elderly patients and those with renal impairment were found to be at highest risk, despite dose-adjustment for creatinine clearance. Since idraparinux was found to be noninferior to warfarin for the composite endpoint of stroke and non-central nervous system (CNS) embolism, a modified compound, biotinylated idraparinux is now being tested in a phase III trial of patients with $CHADS_2$ scores ≥ 2. Biotinylated idraparinux will allow for quick reversal of its anticoagulant effect upon administration of its neutralizing agent, avidin [30].

In contrast to idraparinux, there are multiple oral direct factor Xa inhibitors being developed for stroke prophylaxis

Table 2 Oral anticoagulants currently in phase III trials for stroke prevention in non-valvular AF

Drug	Mechanism of action	Half-life (h)	Bioavailability (%)	Elimination (%)		Dosing	Randomized trial
				Renal	Hepatic		
Dabigatran	Direct thrombin inhibitor	12	4	80	20	Twice daily	RELY [25] 110 mg bid & 150 mg bid vs open label warfarin (goal INR 2–3)
Apixaban	Direct factor Xa inhibitor	12	50	25	75	Twice daily	ARISTOTLE [32] 5 mg bid vs warfarin (goal INR 2–3) double-dummy, double blind
							AVERROES [33] 2.5 mg bid vs aspirin 81–324 mg daily, double-blind
Rivaroxaban	Direct factor Xa inhibitor	5–9	80	33	67	Once daily	ROCKET-AF [35] 20 mg daily vs warfarin (goal INR 2–3) double-dummy, double blind

in AF [31]. These agents are being developed with the goal of no requirement for therapeutic monitoring. The first oral direct factor Xa inhibitor tested in a phase II study was razaxaban. Despite reducing the incidence of DVT with different doses compared with enoxaparin (1.4–8.6% vs. 15.9%), razaxaban was associated with increased bleeding and its development was terminated [32]. Apixaban, a derivative of razaxaban has been shown to decrease the composite endpoint of DVT, pulmonary embolism, and all-cause mortality in patients undergoing total knee replacement without increased bleeding [33]. Apixaban has excellent oral bioavailability (>50%) and is predominantly hepatically cleared (75%, primarily by CYP3A4 and SULT 1A1). Apixaban does not affect the QTc interval and its pharmacokinetics appear to be relatively consistent across age and sex. Currently, apixaban is being compared with warfarin in a randomized, double-blind, event-driven, noninferiority trial that will enroll over 15,000 AF patients (more than one-third will be warfarin naïve) with 1 or more additional risk factors for stroke [34]. Apixaban is also being evaluated against aspirin in a phase III trial, Apixiban vs. aspirin for stroke prevention in atrial fibrillation (AVERROES) [35]. AVERROES will randomize patients who are at moderate risk of stroke ($CHADS_2 = 1$) or intolerant to warfarin.

The other major oral direct factor Xa inhibitor that is currently being evaluated in a phase III study is rivaroxaban. Rivaroxaban is an oxazolidinone derivative with a half-life of 5–9 h. There are no major circulating metabolites of rivaroxaban and its elimination is mixed (one-third renal, two-thirds hepatic). Of all the oral factor Xa inhibitors under development, rivaroxaban has been the most studied, accruing more than 24,000 patients in phase II and phase III trials [36]. Nonetheless, like apixaban, there are no preliminary data for its efficacy in stroke prevention in patients with AF. Rivaroxaban is also being evaluated in a large international, randomized, double-blind, event-driven, noninferiority trial. The Rivaroxaban Once Daily Oral Direct Factor Xa Inhibition Compared with Vitamin K Antagonism for Prevention of Stroke and Embolism Trial in Atrial Fibrillation (ROCKET-AF) is enrolling over 14,000 patients with non-valvular AF who are at high risk of stroke. Candidates for ROCKET-AF must have a history of stroke, TIA, or non-CNS embolism or a $CHADS_2$ score ≥ 3 [37].

While apixaban and rivaroxaban are currently being evaluated as long-term anticoagulants for AF in phase III trials, there are many new oral factor Xa inhibitors in phase I and phase II trials. PRT054021, LY517717, DU-176b, and YM150 are being evaluated for the prevention of venous thromboembolism and slated for testing in patients with AF [31]. Given the track record of prior candidates for alternative oral anticoagulants, hopefully there will be strength in numbers. The goal is to identify a selective anticoagulant with a safety/risk profile that outweighs any similar or improved efficacy. Hopefully, not just one, but many of these compounds will end up as the long-awaited alternative to warfarin.

References

1. Fuster V, Rydén LE, Cannom DS et al (2006) 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 on collaboration with the European Heart Rhythm Association and the Heart Rhythm Society. Circulation 114:e257–e354. doi:10.1161/CIRCULATIONAHA.106.177292
2. Go AS, Hylek EM, Phillips KA et al (2001) 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 285:2370–2375. doi:10.1001/jama.285.18.2370
3. Furberg CD, Psaty BM, Manolio TA et al (1994) Prevalence of atrial fibrillation in elderly subjects (the cardiovascular health study). Am J Cardiol 74:236–241. doi:10.1016/0002-9149(94)90363-8

4. Feinberg WM, Blackshear JL, Laupacis A et al (1995) Prevalence, age distribution, and gender of patients with atrial fibrillation. Analysis and implications. Arch Intern Med 155:469–473. doi:10.1001/archinte.155.5.469
5. Miyasaka Y, Barnes ME, Gersh BJ et al (2006) Secular trends in incidence of atrial fibrillation in Olmsted County, Minnesota, 1980 to 2000, and implications on the projections for future prevalence. Circulation 114:119–125. doi:10.1161/CIRCULATIONAHA.105.595140
6. Benjamin EJ, Wolf PA, D'Agostino RB et al (1998) Impact of atrial fibrillation on the risk of death: the Framingham heart study. Circulation 98:946–952
7. Wolf PA, D'Agostino RB, Belanger AJ et al (1991) Probability of stroke: a risk profile from the Framingham study. Stroke 22:312–318
8. Lairikyengbam SK, Anderson MH, Davies AG (2003) Present treatment options for atrial fibrillation. Postgrad Med J 79:67–73. doi:10.1136/pmj.79.928.67
9. Singer DE, Albers GW, Dalen JE et al (2008) Antithrombotic therapy in atrial fibrillation: American College of chest physicians evidence-based clinical practice guidelines (8th edn). Chest 133:546S–592S. doi:10.1378/chest.08-0678
10. Gage BF, Waterman AD, Shannon W et al (2001) Validation of clinical classification schemes for predicting stroke: results from the national registry of atrial fibrillation. JAMA 285:2864–2870. doi:10.1001/jama.285.22.2864
11. Inoue H, Nozawa T, Hirai T et al (2006) Accumulation of risk factors increases risk of thromboembolic events in patients with nonvalvular atrial fibrillation. Circ J 70:651–656. doi:10.1253/circj.70.651
12. Hylek EM, Go AS, Chang Y et al (2003) Effect of intensity of oral anticoagulation on stroke severity and mortality in atrial fibrillation. N Engl J Med 349:1019–1026. doi:10.1056/NEJMoa022913
13. Garcia DA, Regan S, Henault LE et al (2008) Risk of thromboembolism with short-term interruption of warfarin therapy. Arch Intern Med 168:63–69. doi:10.1001/archinternmed.2007.23
14. Lopes RD, Pieper KS, Horton JR et al (2008) Short- and long-term outcomes following atrial fibrillation in patients with acute coronary syndromes with or without ST-segment elevation. Heart 94:867–873. doi:10.1136/hrt.2007.134486
15. Orford JL, Fasseas P, Melby S et al (2004) Safety and efficacy of aspirin, clopidogrel, and warfarin after coronary stent placement in patients with an indication for anticoagulation. Am Heart J 147:463–467. doi:10.1016/j.ahj.2003.06.004
16. Khurram Z, Chou E, Minutello R et al (2006) Combination therapy with aspirin, clopidogrel and warfarin following coronary stenting is associated with a significant risk of bleeding. J Invasive Cardiol 18:162–164
17. Karjalainen PP, Porela P, Ylitalo A et al (2007) Safety and efficacy of combined antiplatelet-warfarin therapy after coronary stenting. Eur Heart J 28:726–732. doi:10.1093/eurheartj/ehl488
18. Grines CL, Bonow RO, Casey DE Jr et al (2007) Prevention of premature discontinuation of dual antiplatelet therapy in patients with coronary artery stents: a science advisory from the American Heart Association, American College of Cardiology, Society for Cardiovascular Angiography and Interventions, American College of Surgeons, and American Dental Association, with representation from the American College of physicians. J Am Coll Cardiol 49:734–739. doi:10.1016/j.jacc.2007.01.003
19. Becker RC, Meade TW, Berger PB et al (2008) The primary and secondary prevention of coronary artery disease: American College of chest physicians evidence-based clinical practice guidelines (8th edn). Chest 133:776S–814S. doi:10.1378/chest.08-0685
20. Dentali F, Douketis JD, Lim W et al (2007) Combined aspirin-oral anticoagulant therapy compared with oral anticoagulation therapy alone among patients at risk for cardiovascular disease: a meta-analysis of randomized trials. Arch Intern Med 167:117–124. doi:10.1001/archinte.167.2.117
21. Hylek EM, Singer DE (1994) Risk factors for intracranial hemorrhage in outpatients taking warfarin. Ann Intern Med 120:897–902
22. Hylek EM, Skates SJ, Sheehan MA et al (1996) An analysis of the lowest effective intensity of prophylactic anticoagulation for patients with nonrheumatic atrial fibrillation. N Engl J Med 335:540–546. doi:10.1056/NEJM199608223350802
23. Jones M, McEwan P, Morgan CL et al (2005) Evaluation of the pattern of treatment, level of anticoagulation control, and outcome of treatment with warfarin in patients with non-valvular atrial fibrillation: a record linkage study in a large British population. Heart 91:472–477. doi:10.1136/hrt.2004.042465
24. Boos CJ, Lip GY (2006) Ximelagatran: an eulogy. Thromb Res 118:301–304. doi:10.1016/j.thromres.2006.02.012
25. Eriksson BI, Dahl OE, Buller HR et al (2005) A new oral direct thrombin inhibitor, dabigatran etexilate, compared with enoxaparin for prevention of thromboembolic events following total hip or knee replacement: the BISTRO II randomized trial. J Thromb Haemost 3:103–111. doi:10.1111/j.1538-7836.2004.01100.x
26. Stangier J, Rathgen K, Stähle H et al (2007) The pharmacokinetics, pharmacodynamics and tolerability of dabigatran etexilate, a new oral direct thrombin inhibitor, in healthy male subjects. Br J Clin Pharmacol 64:292–303. doi:10.1111/j.1365-2125.2007.02899.x
27. Randomized Evaluation of Long Term Anticoagulant Therapy (RE-LY) With Dabigatran Etexilate NCT00262600. http://clinicaltrials.gov/ct2/show/NCT00262600?term=nct00262600&rank=1. Accessed 01 July 2008
28. Bauer KA (2006) New anticoagulants: anti IIa vs anti Xa–is one better? J Thromb Thrombolysis 21:67–72. doi:10.1007/s11239-006-5579-4
29. Bousser MG, Bouthier J, Buller HR et al (2008) Comparison of idraparinux with vitamin K antagonists for prevention of thromboembolism in patients with atrial fibrillation: a randomised, open-label, non-inferiority trial. Lancet 371:315–321. doi:10.1016/S0140-6736(08)60168-3
30. Evaluation of Weekly Subcutaneous Biotinylated Idraparinux Versus Oral Adjusted-Dose Warfarin to Prevent Stroke and Systemic Thromboembolic Events in Patients with Atrial Fibrillation (BOREALIS-AF) NCT00580216. http://clinicaltrials.gov/ct2/show/NCT00580216?term=nct00580216&rank=1. Accessed 01 July 2008
31. Turpie AG (2007) Direct factor Xa inhibitors in development for the prevention and treatment of thromboembolic diseases. Arterioscler Thromb Vasc Biol 27:1238–1247. doi:10.1161/ATVBAHA.107.139402
32. Kubitza D, Haas S (2006) Novel factor Xa inhibitors for prevention and treatment of thromboembolic diseases. Expert Opin Investig Drugs 15:843–855. doi:10.1517/13543784.15.8.843
33. Lassen MR, Davidson BL, Gallus A et al (2007) The efficacy and safety of apixaban, an oral, direct factor Xa inhibitor, as thromboprophylaxis in patients following total knee replacement. J Thromb Haemost 5:2368–2375. doi:10.1111/j.1538-7836.2007.02764.x
34. Apixaban for the Prevention of Stroke in Subjects with Atrial Fibrillation (ARISTOTLE) NCT00412984. http://clinicaltrials.gov/ct2/show/NCT00412984?term=nct00412984&rank=1. Accessed 01 July 2008
35. Phase A III Study of Apixaban in Patients with Atrial Fibrillation (AVERROES). NCT00496769. http://clinicaltrials.gov/ct2/show/NCT00496769?term=nct00496769&rank=1. Accessed 01 July 2008

36. Piccini JP, Patel MR, Mahaffey KW et al (2008) Rivaroxaban, an oral direct factor Xa inhibitor. Expert Opin Investig Drugs 17:925–937. doi:10.1517/13543784.17.6.925
37. A Study Comparing Once Daily Oral Rivaroxaban with Adjusted-Dose Oral Warfarin for the Prevention of Stroke in Subjects with Non-Valvular Atrial Fibrillation (ROCKET-AF) NCT00403767. http://clinicaltrials.gov/ct2/show/NCT00403767?term=nct00403767&rank=1. Accessed on 01 July 2008

CLINICAL RESEARCH STUDY

THE AMERICAN JOURNAL of MEDICINE®

Sexual Satisfaction and Cardiovascular Disease: The Women's Health Initiative

Jennifer S. McCall-Hosenfeld, MD, MSc,[a,b] Karen M. Freund, MD, MPH,[b] Claudine Legault, PhD,[c] Sarah A. Jaramillo, MS,[c] Barbara B. Cochrane, PhD, RN,[d] JoAnn E. Manson, MD, DrPH,[e] Nanette K. Wenger, MD,[f] Charles B. Eaton, MD, MS,[g] S. Gene McNeeley, MD,[h] Beatriz L. Rodriguez, MD, PhD,[i] Denise Bonds, MD, MPH[j]

[a]VA Boston Healthcare System, Mass; [b]Boston University, Mass; [c]Wake Forest University School of Medicine, Winston-Salem, NC; [d]University of Washington, Seattle; [e]Brigham and Women's Hospital, Boston, Mass; [f]Emory University, Atlanta, Ga; [g]Memorial Hospital of Rhode Island, Pawtucket; [h]Wayne State University, Detroit, Mich; [i]University of Hawaii, Manoa; [j]University of Virginia, Charlottesville.

ABSTRACT

BACKGROUND: Sexual dysfunction in some men is predictive of occult cardiovascular disease. We investigated whether dissatisfaction with sexual activity, a domain of female sexual dysfunction, is associated with prevalent and incident cardiovascular disease in postmenopausal women.

METHODS: Data from the Women′s Health Initiative-Observational Study were used. Subjects who were sexually active in the past year were classified at baseline as sexually satisfied or dissatisfied. We performed multiple logistic regression analyses modeling baseline cardiovascular conditions including myocardial infarction, stroke, coronary revascularization, peripheral arterial disease, congestive heart failure, and angina. We then created Cox proportional hazards models to determine hazard ratios for incident cardiovascular disease by baseline sexual dissatisfaction status.

RESULTS: Dissatisfaction with sexual activity at baseline was significantly associated with prevalent peripheral arterial disease (odds ratio 1.44, 95% confidence interval, 1.15-1.84), but not prevalent myocardial infarction, stroke, coronary revascularization including coronary artery bypass graft and percutaneous transluminal coronary angioplasty, or a composite cardiovascular disease variable. The odds of baseline angina were decreased among those reporting sexual dissatisfaction at baseline (odds ratio 0.77, 95% confidence interval, 0.66-0.86). In both unadjusted and adjusted analyses, dissatisfaction with sexual activity was not significantly related to an increased hazard of any cardiovascular disease.

CONCLUSIONS: Dissatisfaction with sexual activity was modestly associated with an increased prevalence of peripheral arterial disease, even after controlling for smoking status. However, dissatisfaction did not predict incident cardiovascular disease. Although this may represent insensitivity of the sexual satisfaction construct to measure sexual dysfunction in women, it might be due to physiological differences in sexual functioning between men and women.

 • *The American Journal of Medicine (2008) 121, 295-301*

KEYWORDS: Cardiovascular diseases; Cohort studies; Menopause; Physiological; Postmenopause; Sexual dysfunction, Women

Female sexual dysfunction is a prevalent, distressing condition, affecting 43% of the population.[1] Although it has been linked to a higher burden of medical illnesses, specific underlying causes are not well described. One potential candidate is cardiovascular disease. The association of cardiovascular disease with male sexual dysfunction is well documented.[2] Erectile dysfunction is one manifestation of subclinical cardiovascular disease and is a marker for the

Preliminary results from this work were presented at the Society of General Internal Medicine 30th Annual Meeting, Toronto, Ontario, Canada, April 26, 2007.

The Women's Health Initiative program was funded by the National Heart, Lung and Blood Institute of the National Institutes of Health, U.S. Department of Health and Human Services. Dr. McCall-Hosenfeld was supported by a Department of Veterans Affairs Special Fellowship in the Health Issues of Women Veterans.

Requests for reprints should be addressed to Jennifer McCall-Hosenfeld, MD, MSc, Boston University Medical Center, Section of General Internal Medicine, 801 Massachusetts Avenue, Suite 200, Boston, MA 02118.

E-mail address: jmccall@bu.edu

0002-9343/$ -see front matter
doi:10.1016/j.amjmed.2007.11.013

development of subsequent cardiovascular disease in some men.[3]

Basic science and human research suggests that the vascular pathophysiology of male and female sexual dysfunction is similar.[4-6] The regulation of blood flow and clitoral erectile function is governed by the same nitric oxide-cyclic guanosine monophosphate (GMP) pathway in women as erectile function is in men.[4,7] Atherosclerosis of the arterial bed supplying female pelvic anatomy can lead to decreased vaginal engorgement and clitoral erectile insufficiency syndromes,[6] similar to erectile problems in men. Chronic atherosclerotic disease in animal models can cause significant disease of the vagina.[8] To date, the association of cardiovascular disease with female sexual dysfunction has not been explored. Specifically, it has never been shown that subclinical cardiovascular disease impacts female sexual functioning significantly enough to affect sexual satisfaction.

CLINICAL SIGNIFICANCE

- Sexual dysfunction in some men is predictive of cardiovascular disease, but this association has never been examined in women.
- We found no increased prevalence or incidence of cardiovascular disease among sexually active female subjects complaining of dissatisfaction with sexual activity at baseline, over 7.8 years of follow-up.
- Physiological differences between men and women may explain the difference in the role of cardiovascular disease in sexual function.

Current survey instruments[9-11] divide female sexual function into several unique domains, including sexual satisfaction. The International Consensus Development Conference on Female Sexual Dysfunction highlighted personal distress as a necessary element and noted that patient self-report is the preferred method of determining sexual dysfunction in women.[12] Thus, female sexual dysfunction can be measured as self-reported dissatisfaction with sexual activity. Concordantly, sexual satisfaction is an important domain of sexual functioning in women.

The objectives of this study were: to explore the association of sexual satisfaction with prevalent cardiovascular disease and cardiovascular disease risk factors among sexually active postmenopausal women, and to determine if low sexual satisfaction, a domain of female sexual functioning, is predictive of incident cardiovascular disease in women, using the Observational Study cohort of the Women's Health Initiative.

MATERIALS AND METHODS

Subjects

Data from the Women's Health Initiative Observational Study (WHI-OS) were used. Details of recruitment and eligibility screening in the WHI-OS have been described elsewhere.[13] Participants were postmenopausal women aged 50 to 79 years, recruited at 40 clinical centers throughout the United States during 1994 through 1998. There were 93,676 women who participated in the observational cohort[14,15] and were followed for 8-12 years. Institutional review boards approved the study at all participating sites.

Data Collection

Data on sexual satisfaction and baseline cardiovascular disease were collected during one or more initial screening visits. Thereafter, follow-up medical histories of the participants were identified via annual surveys and verification of outcomes. The details of the methodologies used for ascertaining, classifying, and documenting outcomes have been described previously.[16] Each annual follow-up survey with a positive response for cardiovascular disease prompted a review and abstraction of the medical record for that patient. Key outcomes, including cardiovascular disease, were adjudicated by centrally trained WHI physician adjudicators at each clinical center.[16] Of note, follow-up data on sexual satisfaction were not available. Women who indicated at baseline that they did not want to answer the sexual satisfaction question, and those with missing data for this item were excluded from analyses. To isolate medical rather than social etiologies of sexual satisfaction, subjects who reported no sexual activity with a partner in the last year also were excluded.

Definition of Variables

Satisfaction with current sexual activity was assessed by a single survey item categorized along a 4-point Likert-type scale, "How satisfied are you with your current sexual activities, either with a partner or alone?" (1 = very unsatisfied, 2 = a little unsatisfied, 3 = somewhat satisfied, 4 = very satisfied). After qualitatively assessing the response variables to ensure similar direction between extreme variables and nearest middle range response, the responses were dichotomized into satisfied (responses 3 and 4) versus unsatisfied (responses 1 and 2) to simplify analysis.

The presence of cardiovascular disease at baseline was defined as a self-reported history of acute myocardial infarction, stroke, or coronary revascularization procedure (coronary artery bypass graft or percutaneous transluminal coronary angioplasty). Related cardiovascular problems, including congestive heart failure, peripheral arterial disease, and angina also were examined.

Covariates included cardiovascular disease risk factors and mitigants including physical activity, smoking status, hypertension, family history of myocardial infarction, high

cholesterol requiring pills, beta-blocker use, diabetes, and body mass index (BMI).

Covariates included health-related factors that might affect sexual functioning. Overall health-related quality of life was measured using the general health subscale of the SF-36 (scores range from 0-100 with a higher score indicating a more favorable health state).[17] Current depressive symptoms were measured using an 8-item scale designed to screen for depressive disorders,[18] composed of 2 items from the Center for Epidemiologic Studies Depression Scale[19] and 2 items from the Diagnostic Interview Schedule.[20] Higher scores on this scale indicate more severe depression. Other health-related conditions included as covariates in models were a history of cervical, ovarian, or endometrial cancer, a history of hysterectomy, and parity, all measured by patient self-report at baseline. Medications including oral contraceptive use ever, hormone replacement therapy use at baseline, and selective serotonin reuptake inhibitor (SSRI) use also were examined.

Demographic covariates included race and ethnicity, marital status, family income, education, employment status, and sexual orientation, all determined by patient self-report from the baseline questionnaires.

For prospective analyses, fatal and nonfatal incident cardiovascular disease in the WHI-OS cohort were defined as incident acute myocardial infarction, coronary death, stroke, or coronary revascularization that occurred during follow-up. Related incident diseases, including congestive heart failure, peripheral arterial disease, and angina also were examined. Subjects who reported prevalent cardiovascular disease or related illnesses at baseline were excluded from analyses of prospective data.

Statistical Analysis

To determine the association between sexual satisfaction and baseline cardiovascular conditions, logistic regression models examining each of the baseline cardiovascular conditions (myocardial infarction, stroke, coronary revascularization, and a composite endpoint including each of the above, peripheral arterial disease, angina, and congestive heart failure) were developed. These models included demographic variables, baseline medical conditions, and cardiovascular disease risk factors that were significantly associated with sexual satisfaction or cardiovascular disease in bivariate analyses. Backwards and stepwise selection procedures were utilized, with both 0.05 and 0.20 selected as entry and retention criteria. With backwards selection, the initial model includes all variables, which are then deleted from the model one at a time until all the variables remaining in the model meet the retention criteria. With stepwise selection, variables are entered one by one to the model based on the entry criterion. At each step, any of the variables in the model that does not meet the retention criterion is deleted. The lists of variables in the final models did not change under these different assumptions. After creating these final models for each cardiovascular outcome, sexual satisfaction status was included in the models.

To determine whether sexual satisfaction at baseline was associated with incident cardiovascular disease, variables significant at the 0.05 level in bivariate analyses, as well as clinically important variables related to cardiovascular disease, were fitted into Cox proportional hazards model, with time to incident cardiovascular disease as the response variable. As in the prevalence analysis, variables were retained in the models if they met a prespecified retention criterion of 0.05. Hazard ratios and their respective confidence intervals are reported.

RESULTS

Baseline characteristics of the WHI-OS participants have been described elsewhere.[13] Fifty-two percent (48,300) of these women reported that they had been sexually active with another person in the past year. Of these, 96% (46,525) answered the sexual satisfaction question. Overall, 77% (35,719) of respondents reported satisfaction with sexual activity.

Women who reported that they were satisfied sexually were older and had higher family incomes than women who reported that they were not satisfied sexually (Table 1). Sexual satisfaction also was associated with better overall physical health and fewer current depressive symptoms.

Satisfaction with sexual activity, compared with dissatisfaction, was modestly associated with several cardiovascular disease risk factors, including more physical activity, never smoking, and normal BMI (Table 2). A slightly higher proportion of subjects on beta-blockers at baseline reported dissatisfaction with sexual activity, but this difference was minimal. No association was found between sexual satisfaction and other cardiovascular disease risk factors, including hypertension, menopausal hormone therapy, diabetes, family history of myocardial infarction, and hyperlipidemia.

In logistic regression modeling, dissatisfaction with sexual activity was significantly associated with prevalent peripheral arterial disease (odds ratio 1.44, 95% confidence interval, 1.15-1.82; Table 3). The odds of prevalent angina were lower among those reporting sexual dissatisfaction (odds ratio 0.77, 95% confidence interval, 0.66-0.90, $P < 0.001$). Sexual dissatisfaction at baseline was not significantly associated with prevalent myocardial infarction, stroke, coronary revascularization, the composite cardiovascular disease variable, or congestive heart failure.

To determine if women with cardiovascular disease abstained from sexual activity, an interaction term between levels of reported sexual activity in the past year and sexual satisfaction was included in each of the models. Level of sexual activity in the past year was defined as any versus none. If significant, the interaction term would suggest that the association between sexual satisfaction and cardiovascular outcome differed between those with reported sexual activity and those with no reported sexual activity in the past year. This interaction term was not significant in any model, thus, exclusion of those who were not sexually active due to cardiovascular disease was unlikely to explain our negative findings.

Table 1 Baseline Demographics by Sexual Satisfaction Status (Women's Health Initiative Observational Study)

	Satisfied Sexually (n = 35,719)		Not Satisfied Sexually (n = 10,806)	
	n	%	n	%
Categorical variables				
Age (years)*				
50-59	14,741	41%	4906	45%
60-69	15,673	44%	4513	42%
70-79	5305	15%	1387	13%
Race/ethnicity*				
White	30,629	86%	9325	86%
Black	2269	6%	776	7%
Hispanic	1252	4%	385	4%
Other	1490	4%	299	3%
Marital status*				
Never married	345	1%	161	1%
Divorced/separated	2632	7%	1295	12%
Widowed	1508	4%	580	5%
Married/partnered	31,083	87%	8720	81%
Family income*				
<$10,000-19,999	2321	7%	874	9%
$20,000-$49,999	12,571	38%	3992	39%
$50,000-74,999	8221	25%	2476	24%
$75,000+	10,191	31%	2782	27%
Measurement variables	Mean	(SD)	Mean	(SD)
Depression score*†	0.03	0.10	0.06	0.15
Overall health*‡	76	17	74	19

*P <.001.
†Scale: −8.2-4.0. Higher scores indicate more severe depressive symptoms.
‡Scale: 0-100. Higher scores indicate better general health.

To determine whether our results were confounded by the exclusion of women who abstained from sexual activity because of cardiovascular disease and its symptoms and were dissatisfied because of this, we assessed whether subjects worried that sexual activity would affect their health. Worry that sexual activity would affect their health showed little difference between sexually satisfied and dissatisfied subjects. Similar percentages of both the satisfied (99%) and dissatisfied (96%) respondents were either not at all worried, or only a little worried. Thus, dissatisfaction with sexual activity due to fear of cardiovascular health-related consequences is an unlikely explanation for these findings.

To determine if inclusion of cardiovascular risk factors over-adjusted the regression models, a separate series of regressions models excluding the cardiovascular disease risk factors were performed and no significant changes in overall results were found.

For our prospective analyses, subjects were followed for an average of 7.8 ± 1.4 years. In unadjusted and adjusted analyses of follow-up data, no increased incidence or hazard of cardiovascular disease among sexually satisfied versus dissatisfied participants was found (Table 4). No changes in these results were found when cardiovascular risk factors were excluded from the models.

DISCUSSION

The present study identified a higher prevalence of peripheral arterial disease among women who reported sexual dissatisfaction at baseline. However, no association between baseline sexual satisfaction and the other cardiovascular diseases investigated, including myocardial infarction, stroke, coronary revascularization, and congestive heart failure was found. Decreased sexual satisfaction at baseline did not predict incident cardiovascular disease.

A number of factors may account for our overall null findings. Sexual satisfaction is a complex, multi-factorial construct. Psychosocial stressors, comorbid medical conditions, and nonvascular organic pathology play a large role. Although we designed our study to explore and control for other potential contributors to sexual satisfaction in these women, including medical, psychiatric, and social covariates, it is likely that residual confounding remained.

Table 2 Baseline Prevalence of Cardiovascular Disease Risk Factors by Sexual Satisfaction Status (Women's Health Initiative Observational Study)

CVD Risk Factor	Satisfied Sexually (n = 35,719)		Not Satisfied Sexually (n = 10,806)	
Exercise*				
Little or no activity	16,242	46%	5134	48%
2 or more times per week	19,167	54%	5595	52%
Smoking*				
Never smoked	18,897	53%	5208	49%
Past smoker	14,616	41%	4813	45%
Current smoker	1825	5%	654	6%
Body mass index*				
<18.5 Underweight	349	1%	122	1%
18.5-24.9 Normal	15,583	37%	4565	35%
25.0-29.9 Overweight	12,096	28%	3657	28%
≥30 Obesity	7301	17%	2343	18%
Hypertension (P = .31)				
Never hypertensive	25,007	71%	7522	71%
Treated hypertensive	7697	22%	2331	22%
Untreated hypertensive	2497	7%	794	7%
Baseline hormone replacement therapy use (P = .13)	15,572	44%	4800	44%
Diabetes (P = .17)	1470	4%	477	4%
Family history of myocardial infarction (P = .19)	4821	28%	1533	29%
Treated hyperlipidemia (P = .12)	4513	13%	1427	13%
Beta-blocker use†	1602	4.5%	533	5.0%

*P <.001.
†P <.01.

Sexual satisfaction may have a different meaning for women than it has for men. The underlying mechanism for sexual functioning may differ between men and women, with cardiovascular risk factors playing a stronger role in sexual functioning among men and other variables of greater importance in female sexual functioning. Although data on sexual satisfaction in men are limited, previous studies have shown a strong association between sexual

Table 3 Odds of Prevalent Cardiovascular Disease by Sexual Satisfaction Status at Baseline (Women's Health Initiative Observational Study)

Baseline Cardiovascular Disease	OR (95% CI)	aOR (95% CI)
Myocardial infarction*	1.11 (0.94-1.31)	1.09 (0.88-1.36)
Stroke†	1.20 (0.98-1.47)	1.23 (0.99-1.52)
Coronary revascularization‡	0.89 (0.73-1.08)	0.92 (0.72-1.17)
Composite cardiovascular disease§	1.04 (0.92-1.17)	0.94 (0.78-1.11)
Peripheral arterial disease‖	1.52 (1.30-1.79)	1.44 (1.15-1.82)
Angina¶	0.98 (0.88-1.09)	0.77 (0.66-0.90)
Congestive heart failure**	0.95 (0.71-1.28)	0.93 (0.63-1.36)

OR = odds ratio; CI = confidence interval; aOR = adjusted odds ratio. Ref = satisfied.

All adjusted models included the following variables: demographics (race/ethnicity, marital status, family income, education, employment, sexual orientation); medical factors (general health, depressive symptoms, history of cervical, ovarian, or endometrial cancer, history of hysterectomy, and oral contraceptive [OC] use ever, selective serotonin reuptake inhibitor [SSRI] use); cardiovascular risk factors (family history of myocardial infarction, exercise, body mass index [BMI], hormone replacement therapy [HRT] use, diabetes, hyperlipidemia, beta-blocker use). Variables removed from models if $P < .05$.

*Myocardial infarction model variables removed: race/ethnicity, marital status, family income, sexual orientation, history of cervical, ovarian, or endometrial cancer, history of hysterectomy, OC use ever, baseline HRT usage, SSRI use, exercise, BMI.

†Stroke model variables removed: race/ethnicity, marital status, family income, education, sexual orientation, history of cervical, ovarian, or endometrial cancer, parity, OC use ever, baseline HRT usage, SSRI use, exercise, family history of myocardial infarction, BMI.

‡Coronary revascularization model variables removed: race/ethnicity, marital status, family income, employment status, sexual orientation, depression score (continuous), history of cervical, ovarian, or endometrial cancer, history of hysterectomy, parity, OC use ever, baseline HRT usage, SSRI use, exercise, BMI.

§Composite cardiovascular disease model variables removed: race/ethnicity, marital status, family income, sexual orientation, history of cervical, ovarian, or endometrial cancer, parity, OC use ever, baseline HRT usage, SSRI use, exercise, BMI.

‖Peripheral arterial disease model variables removed: race/ethnicity, marital status, family income, employment status, sexual orientation, history of cervical, ovarian, or endometrial cancer, OC use, exercise.

¶Angina model variables removed: race/ethnicity, marital status, family income, employment status, history of cervical, ovarian, or endometrial cancer, parity, OC use ever, exercise.

**CHF model variables removed: race/ethnicity, marital status, family income, education, sexual orientation, history of cervical, ovarian, or endometrial cancer, history of hysterectomy, parity, OC use ever, baseline HRT usage, SSRI use, BMI.

Table 4 Hazards of Incident Cardiovascular Disease by Sexual Satisfaction Status at Baseline (Women's Health Initiative Observational Study)

CVD Outcome	HR (95% CI)	aHR (95% CI)
Total cardiovascular disease (fatal and nonfatal)*	1.07 (0.95-1.20)	1.12 (0.95-1.31)
Myocardial infarction†	1.11 (0.90-1.38)	1.08 (0.87-1.35)
Coronary revascularization‡	1.12 (0.95-1.32)	1.19 (0.96-1.48)
Stroke§	1.06 (0.86-1.31)	0.99 (0.80-1.23)
Peripheral arterial disease‖	0.83 (0.53-1.29)	0.79 (0.51-1.25)
Angina¶	0.94 (0.79-1.11)	0.94 (0.76-1.16)
Congestive heart failure**	1.08 (0.87-1.35)	1.18 (0.93-1.49)
Carotid artery disease††	0.87 (0.60-1.25)	0.92 (0.63-1.36)

Ref = satisfied.

All adjusted models included the following variables: demographics (race/ethnicity, marital status, family income, education, employment, sexual orientation); medical factors (general health, depressive symptoms, history of cervical, ovarian, or endometrial cancer, history of hysterectomy, and oral contraceptive use ever, selective serotonin reuptake inhibitor [SSRI] use); cardiovascular risk factors (family history of myocardial infarction, exercise, body mass index [BMI], hormone replacement therapy [HRT] use, diabetes, hyperlipidemia, beta-blocker use). Variables removed from models if $P < .05$.

*Total cardiovascular disease model variables removed: race/ethnicity, marital status, education, employment status, sexual orientation, depression score, history of cervical, ovarian, or endometrial cancer, history of hysterectomy, parity, oral contraceptive (OC) use ever, baseline HRT use, SSRI use, exercise, BMI.

†Myocardial infarction model variables removed: marital status, education, employment status, sexual orientation, depression score, history of cervical, ovarian, or endometrial cancer, history of hysterectomy, OC use ever, baseline HRT use, SSRI use, exercise, family history of myocardial infarction, BMI.

‡Coronary revascularization model variables removed: race/ethnicity, marital status, employment status, sexual orientation, depression score, history of cervical, ovarian, or endometrial cancer, history of hysterectomy, OC use ever, baseline HRT use, SSRI use, exercise, BMI.

§Stroke model variables removed: race/ethnicity, marital status, family income, education, employment status, sexual orientation, depression score, SF-36 subscale (overall health), history of hysterectomy, parity, OC use ever, baseline HRT use, SSRI use, exercise, family history of myocardial infarction, BMI.

‖Peripheral arterial disease model variables removed: race/ethnicity, marital status, family income, employment status, sexual orientation, depression score, SF-36 subscale (overall health), history of cervical, ovarian, or endometrial cancer, history of hysterectomy, parity, OC use ever, baseline HRT use, SSRI use, exercise, family history of myocardial infarction, diabetes, BMI.

¶Angina model variables removed: marital status, education, employment status, sexual orientation, depression score, history of cervical, ovarian, or endometrial cancer, parity, OC use ever, baseline HRT use, SSRI use, exercise, smoking, BMI.

**Congestive heart failure model variables removed: marital status, education, employment status, sexual orientation, depression score, history of hysterectomy, parity, OC use ever, baseline HRT use, SSRI use, exercise, family history of myocardial infarction.

††Carotid artery disease model variables removed: race/ethnicity, marital status, family income, education, employment status, sexual orientation, depression score, history of cervical, ovarian, or endometrial cancer, history of hysterectomy, parity, OC use ever, baseline HRT use, SSRI use, exercise, family history of myocardial infarction.

dysfunction and cardiovascular disease or cardiovascular risk factors in men, including ischemic heart disease, hypertension, diabetes mellitus, and smoking.[2,21,22] Recently, erectile dysfunction has been implicated as a sentinel sign of cardiovascular disease in men.[3]

Few studies have examined this association in women. Previous work has shown that among type I diabetics, diabetes control measured by factors such as A1C and BMI was associated with sexual dysfunction in men, but not in women, whereas psychological variables were prominent for both sexes.[23] Additionally, the utility of phophodiesterase inhibitors has been shown to differ between men and women. Although widely used to treat sexual dysfunction in men, sildenafil has shown mixed results when used for treatment of sexual dysfunction in women,[5,24] suggesting a more complex mechanism of sexual functioning in women.

We found increased prevalence, but not incidence, of peripheral arterial disease among women reporting low sexual satisfaction, even after controlling for smoking. One possible explanation for this disparity is that different adjudication procedures were used for prevalent versus incident disease. Prevalent peripheral arterial disease was determined based on patient self-report, whereas incident disease was formally adjudicated by trained physicians using standardized criteria. If, at baseline, subjects tended to over-report peripheral arterial disease, this may partially explain these findings. Alternatively, there may be an association between peripheral arterial disease and decreased sexual satisfaction at baseline because peripheral arterial disease predicts decreased sexual satisfaction, rather than the reverse. Because we do not have follow-up data on sexual satisfaction, we are unable to explore this further.

There are several strengths of this study. The WHI-OS is a large cohort study with good representation of women across the country and across racial and ethnic groups. Rigorous adjudication procedures were used to determine incident cardiovascular disease. Additionally, this is the first study, to our knowledge, to describe the association of sexual satisfaction with either prevalent or incident cardiovascular disease in women.

An important limitation of this work is that the sexual satisfaction construct has not been formally validated. The survey question used to determine sexual satisfaction has strong face validity and has similar wording to questions on other validated instruments. However, we cannot fully appreciate the extent to which this question measures the construct of sexual satisfaction versus other constructs that may be highly related to sexual satisfaction, such as overall life satisfaction.

CONCLUSIONS

Sexual dysfunction is a prevalent condition in postmenopausal women, as is cardiovascular disease. Many of the same pathophysiological mechanisms known to be risk factors for cardiovascular disease have been proposed to be responsible for sexual dysfunction in postmenopausal women. However, in our population of sexually active postmenopausal women, dissatisfaction with sexual activity was not predictive of incident of cardiovascular disease. This may be due to physiological differences in sexual functioning between men and women. Further study may better elucidate whether cardiovascular disease is an important element of female sexual functioning.

ACKNOWLEDGMENTS

The WHI program is funded by the National Heart, Lung and Blood Institute, US Department of Health and Human Services. The funding organization had representation on the steering committee, which governed the design and conduct of the study, the interpretation of the data, and the preparation and approval of manuscript. The National Heart, Lung, and Blood Institute Program Office reviewed the manuscript before publication.

WHI Investigators by Clinical Center

Program Office: (National Heart, Lung, and Blood Institute, Bethesda, Md) Elizabeth Nabel, Jacques Rossouw, Shari Ludlam, Linda Pottern, Joan McGowan, Leslie Ford, and Nancy Geller.

Clinical Coordinating Center: (Fred Hutchinson Cancer Research Center, Seattle, Wash) Ross Prentice, Garnet Anderson, Andrea LaCroix, Charles L. Kooperberg, Ruth E. Patterson, Anne McTiernan; (Wake Forest University School of Medicine, Winston-Salem, NC) Sally Shumaker; (Medical Research Labs, Highland Heights, Ky) Evan Stein; (University of California at San Francisco, San Francisco, Calif) Steven Cummings.

Clinical Centers: (Albert Einstein College of Medicine, Bronx, NY) Sylvia Wassertheil-Smoller; (Baylor College of Medicine, Houston, Tex) Aleksandar Rajkovic; (Brigham and Women's Hospital, Harvard Medical School, Boston, Mass) JoAnn Manson; (Brown University, Providence, RI) Annlouise R. Assaf; (Emory University, Atlanta, Ga) Lawrence Phillips; (Fred Hutchinson Cancer Research Center, Seattle, Wash) Shirley Beresford; (George Washington University Medical Center, Washington, DC) Judith Hsia; (Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance, Calif) Rowan Chlebowski; (Kaiser Permanente Center for Health Research, Portland, Ore) Evelyn Whitlock; (Kaiser Permanente Division of Research, Oakland, Calif) Bette Caan; (Medical College of Wisconsin, Milwaukee, Wis) Jane Morley Kotchen; (MedStar Research Institute/Howard University, Washington, DC) Barbara V. Howard; (Northwestern University, Chicago/Evanston, Ill) Linda Van Horn; (Rush Medical Center, Chicago, Ill) Henry Black; (Stanford Prevention Research Center, Stanford, Calif) Marcia L. Stefanick; (State University of New York at Stony Brook, Stony Brook, NY) Dorothy Lane; (The Ohio State University, Columbus, Ohio) Rebecca Jackson; (University of Alabama at Birmingham,

Birmingham, Ala) Cora E. Lewis; (University of Arizona, Tucson/Phoenix, Ariz) Tamsen Bassford; (University at Buffalo, Buffalo, NY) Jean Wactawski-Wende; (University of California at Davis, Sacramento, Calif) John Robbins; (University of California at Irvine, Calif) F. Allan Hubbell; (University of California at Los Angeles, Los Angeles, Calif) Howard Judd; (University of California at San Diego, LaJolla/Chula Vista, Calif) Robert D. Langer; (University of Cincinnati, Cincinnati, Ohio) Margery Gass; (University of Florida, Gainesville/Jacksonville, Fla) Marian Limacher; (University of Hawaii, Honolulu, Hawaii) David Curb; (University of Iowa, Iowa City/Davenport, Iowa) Robert Wallace; (University of Massachusetts/Fallon Clinic, Worcester, Mass) Judith Ockene; (University of Medicine and Dentistry of New Jersey, Newark, NJ) Norman Lasser; (University of Miami, Miami, Fla) Mary Jo O'Sullivan; (University of Minnesota, Minneapolis, Minn) Karen Margolis; (University of Nevada, Reno, Nev) Robert Brunner; (University of North Carolina, Chapel Hill, NC) Gerardo Heiss; (University of Pittsburgh, Pittsburgh, Pa) Lewis Kuller; (University of Tennessee, Memphis, Tenn) Karen C. Johnson; (University of Texas Health Science Center, San Antonio, Tex) Robert Brzyski; (University of Wisconsin, Madison, Wis) Gloria E. Sarto; (Wake Forest University School of Medicine, Winston-Salem, NC) Denise Bonds; (Wayne State University School of Medicine/Hutzel Hospital, Detroit, Mich) Susan Hendrix.

References

1. Laumann EO, Paik A, Rosen RC. Sexual dysfunction in the United States: prevalence and predictors. *JAMA.* 1999;281(6):537-544.
2. Hafez ESE, Hafez SD. Erectile dysfunction: anatomical parameters, etiology, diagnosis, and therapy. *Arch Androl.* 2005;51(1):15-31.
3. Thompson IM, Tangen CM, Goodman PJ, et al. Erectile dysfunction and subsequent cardiovascular disease. *JAMA.* 2005;294(23):2996-3002.
4. Gragasin FS, Michelakis ED, Hogan A, et al. The neurovascular mechanism of clitoral erection: nitric oxide and cGMP-stimulated activation of BKCa channels. *FASEB J.* 2004;18(12):1382-1391.
5. Berman JR, Berman LA, Toler SM, et al; Sildenafil Study Group. Safety and efficacy of sildenafil citrate for the treatment of female sexual arousal disorder: a double-blind, placebo controlled study. *J Urol.* 2003;170(6 Pt 1):2333-2338.
6. Goldstein I, Berman JR. Vasculogenic female sexual dysfunction: vaginal engorgement and clitoral erectile insufficiency syndromes. *Int J Impot Res.* 1998;10(Suppl 2):S84-S90.
7. Kim SW, Jeong SJ, Munarriz R, et al. Role of the nitric oxide-cyclic GMP pathway in regulation of vaginal blood flow. *Int J Impot Res.* 2003;15(5):355-361.
8. Park K, Tarcan T, Goldstein I, et al. Atherosclerosis-induced chronic arterial insufficiency causes clitoral cavernosal fibrosis in the rabbit. *Int J Impot Res.* 2000;12(2):111-116.
9. Rosen R, Brown C, Heiman J, et al. The Female Sexual Function Index (FSFI): a multidimensional self-report instrument for the assessment of female sexual function. *J Sex Marital Ther.* 2000;26(2):191-208.
10. Rosen RC, Lobo RA, Block BA, et al. Menopausal Sexual Interest Questionnaire (MSIQ): a unidimensional scale for the assessment of sexual interest in postmenopausal women. *J Sex Marital Ther.* 2004; 30(4):235-250.
11. Rust J, Golombok S. The GRISS: a psychometric instrument for the assessment of sexual dysfunction. *Arch Sex Behav.* 1986;15(2):157-165.
12. Basson R, Berman J, Burnett A, et al. Report of the international consensus development conference on female sexual dysfunction: definitions and classifications. *J Urol.* 2000;163(3):888-893.
13. Langer RD, White E, Lewis CE, et al. The Women's Health Initiative Observational Study: baseline characteristics of participants and reliability of baseline measures. *Ann Epidemiol.* 2003;13(9 Suppl):S107-S121.
14. Howard BV, Kuller L, Langer R, et al. Risk of cardiovascular disease by hysterectomy status, with and without oophorectomy: the Women's Health Initiative Observational Study. *Circulation.* 2005;111(12): 1462-1470.
15. Hsia J, Aragaki A, Bloch M, et al; WHI Investigators. Predictors of angina pectoris versus myocardial infarction from the Women's Health Initiative Observational Study. *Am J Cardiol.* 2004;93(6):673-678.
16. Curb JD, McTiernan A, Heckbert SR, et al. Outcomes ascertainment and adjudication methods in the Women's Health Initiative. *Ann Epidemiol.* 2003;13(9 Suppl):S122-S128.
17. Hays R, Sherbourne C, Mazel R. The RAND 36-Item Health Survey 1.0. *Health Econ.* 1993;2(3):213-215.
18. Burnam M, Wells K, Leake B, Lansverk J. Development of a brief screening instrument for detecting depressive disorders. *Med Care.* 1988;26:775-789.
19. Radloff L. The CES-D scale: a self-report depression scale for research in the general population. *Appl Psychol Meas.* 1977;1:385-401.
20. Robins LN, Helzer JE, Croughan J, Ratcliff KS. National Institute of Mental Health Diagnostic Interview Schedule: its history, characteristics, and validity. *Arch Gen Psychiatry.* 1981;38:381-389.
21. Feldman HA, Johannes CB, Derby CA, et al. Erectile dysfunction and coronary risk factors: prospective results from the Massachusetts male aging study. *Prev Med.* 2000;30(4):328-338.
22. Rendell MS, Rajfer J, Wicker PA, Smith MD. Sildenafil for treatment of erectile dysfunction in men with diabetes: a randomized controlled trial. *JAMA.* 1999;281(5):421-426.
23. Enzlin P, Mathieu C, Van Den Bruel A, et al. Prevalence and predictors of sexual dysfunction in patients with type 1 diabetes. *Diabetes Care.* 2003;26(2):409-414.
24. Basson R, McInnes R, Smith MD, et al. Efficacy and safety of sildenafil citrate in women with sexual dysfunction associated with female sexual arousal disorder. *J Womens Health Gend Based Med.* 2002; 11(4):367-377.

Correlates of Sexual Satisfaction Among Sexually Active Postmenopausal Women in the Women's Health Initiative-Observational Study

Jennifer S. McCall-Hosenfeld, MD, MSc[1], Sarah A. Jaramillo, MS[2], Claudine Legault, PhD[2], Karen M. Freund, MD, MPH[3], Barbara B. Cochrane, PhD, RN[4], JoAnn E. Manson, MD, DrPH[5], Nanette K. Wenger, MD[6], Charles B. Eaton, MD, MS[7], Beatriz L. Rodriguez, MD, PhD[8], S. Gene McNeeley, MD[9], and Denise Bonds, MD, MPH[10]

[1]Division of General Internal Medicine and Department of Public Health Sciences, Pennsylvania State University College of Medicine, Hershey, PA, USA; [2]Division of Public Health Sciences, Wake Forest University School of Medicine, Winston-Salem, NC, USA; [3]Women's Health Unit, Department of Medicine and Women's Health Interdisciplinary Research Center, Boston University School of Medicine, Boston, MA, USA; [4]School of Nursing, University of Washington, Seattle, WA, USA; [5]Division of Preventive Medicine, Brigham and Women's Hospital Harvard Medical School, Boston, MA, USA; [6]Division of Cardiology, Emory University School of Medicine, Atlanta, GA, USA; [7]Brown University, Memorial Hospital of Rhode Island, Pawtucket, RI, USA; [8]Departments of Geriatric Medicine, Public Health Sciences and Epidemiology, University of Hawaii, Manoa, HI, USA; [9]Hutzel Women's Health, Wayne State University, Detroit, MI, USA; [10]Departments of Public Health Sciences and Medicine, University of Virginia, Charlottesville, VA, USA.

BACKGROUND: Satisfaction with sexual activity is important for health-related quality of life, but little is known about the sexual health of postmenopausal women.

OBJECTIVE: Describe factors associated with sexual satisfaction among sexually active postmenopausal women.

DESIGN: Cross-sectional analysis.

PARTICIPANTS: All members of the Women's Health Initiative-Observational Study (WHI-OS), ages 50–79, excluding women who did not respond to the sexual satisfaction question or reported no partnered sexual activity in the past year (N=46,525).

MEASUREMENTS: Primary outcome: dichotomous response to the question, "How satisfied are you with your sexual activity (satisfied versus unsatisfied)?" Covariates included sociodemographic factors, measures of physical and mental health, and gynecological variables, medications, and health behaviors related to female sexual health.

RESULTS: Of the cohort, 52% reported sexual activity with a partner in the past year, and 96% of these answered the sexual satisfaction question. Nonmodifiable factors associated with sexual dissatisfaction included age, identification with certain racial or ethnic groups, marital status, parity, and smoking history. Potentially modifiable factors included lower mental health status and use of SSRIs. The final model yielded a c-statistic of 0.613, reflecting only a modest ability to discriminate between the sexually satisfied and dissatisfied.

CONCLUSIONS: Among postmenopausal women, the variables selected for examination yielded modest ability to discriminate between sexually satisfied and dissatisfied participants. Further study is necessary to better describe the cofactors associated with sexual satisfaction in postmenopausal women.

KEY WORDS: sexual dysfunction; physiological; sexual dysfunctions; psychological; women; menopause; postmenopause; cohort studies.
J Gen Intern Med 23(12):2000–9
DOI: 10.1007/s11606-008-0820-9

Preliminary results from this work were presented at the Society of General Internal Medicine 30th Annual Meeting, Toronto, Ontario, Canada, on April 26, 2007.
Received March 27, 2008
Revised September 16, 2008
Accepted September 18, 2008
Published online October 7, 2008

BACKGROUND

Sexual satisfaction is a domain of female sexual functioning measured by validated instruments.[1–3] Women remain sexually active throughout life,[4,5] and female sexual dysfunction is prevalent.[6] Older women have increased risk for sexual dysfunction,[6] which may manifest as decreased sexual satisfaction. Medical and psychiatric illness, lower socioeconomic status,[6,7] postmenopausal hormonal changes,[5,8,9] and pelvic surgery[10,11] may contribute. Premenopausal sexual desire and response disorders can worsen after menopause.[12]

Sexual satisfaction among postmenopausal women is inadequately described. Cross-sectional studies of female sexual functioning are compromised by low response rates[13,14] and use of convenience samples.[14] Large surveys[6,15] may select participants with nonrepresentative sexual attitudes. Prevalence estimates vary widely across studies.[14]

In the current project, cross-sectional data collected from the Women's Health Initiative Observational Cohort (WHI-OS[16,17]) were used to describe the prevalence and correlates of low sexual satisfaction in postmenopausal women.

METHODS

Inclusion Criteria. Data from 93,676 participants were collected at baseline during screening visits. For most participants, a single screening visit was sufficient prior to data collection. For a minority of participants, up to three screening visits were required. Participants were postmenopausal women aged 50 to 79 years at the time of enrollment, recruited at 40 different clinical centers throughout the United States during 1994 through 1998.[18] Institutional review boards at all sites approved the study; informed consent was obtained from all participants. Liberal inclusion criteria enhanced generalizability. Inclusion criteria were postmenopausal status, intention to reside in the area for at least 3 years, and ability to provide written informed consent.[18,19]

Exclusion Criteria. Participants were excluded from the WHI if they had comorbid conditions that limited their survival to less than 3 years or if they had conditions (such as substance abuse or dementia) that would affect their ability to follow-up.[18,19] Because lack of an available partner may contribute to low sexual satisfaction in older women,[6] participants who answered "no" to the question "Did you have sexual activity with a partner in the past year?" were excluded from this analysis.

Definitions of Variables. Data were obtained from survey responses as previously described.[17,20,21] Sexual satisfaction was assessed by a single item categorized along a four-point Likert-type scale from very unsatisfied to very satisfied, "How satisfied are you with your current sexual activities, either with a partner or alone?"[22,23] Satisfaction with sexual frequency was assessed by the question, "Are you satisfied with the frequency of your sexual activity, or would you like to have sex more or less often?" Response categories included "less often," "satisfied with current frequency," "more often," and "don't want to answer."

Variables chosen for examination were the result of a comprehensive literature review of factors previously reported to be associated with female sexual dysfunction. A conceptual model for correlates of sexual satisfaction included the following interdependent categories of variables: socioeconomic/demographic variables, physical health status, mental health status, medications, gynecological history, and health behaviors. By consensus and discussion among coauthors, variables that were more distal along our proposed causal pathway were removed. For example, we excluded specific disease states that may impact sexual functioning in favor of a more proximal measure of the effect of these diseases, the subscales of the RAND 36-Item Health Survey (SF-36) that relate to physical health.

Socioeconomic/demographic variables selected for examination were age,[5,6,24,25] marital status,[6,24] family income,[6] race/ethnicity,[6,15,26] education level,[6,26] and type of employment.[6,26] Sexual orientation[27,28] was defined as described by Valanis and colleagues.[21]

For physical health,[6,25,26,29–31] the four subscales of the SF-36 that aggregate to physical health (physical functioning, role limitations due to physical problems, bodily pain, and general health) were assessed.[32,33] Mental health[6,34–36] was assessed using the four subscales of the SF-36 that aggregate to mental health (vitality, social functioning, role limitations due to emotional problems, and mental health). These validated scales are scored from 0–100, with higher scores indicating better health-related quality of life.

Gynecological histories examined included previous oral contraceptive use,[37] a history of gynecological cancer,[38] a history of hysterectomy,[39] and increasing parity.[40] Medications selected for examination included selective serotonin-reuptake inhibitor (SSRI)[41,42] use, determined from a pharmacy database [Master Drug Data Base (MDBB): Medi-Span, Indianapolis, IN],[43,44] and hormone therapy use.[45] Health behaviors, including exercise,[46] smoking,[24] body mass index (BMI),[26,47] and alcohol use[48,49], were also assessed.

Table 1. Demographics by Sexual Satisfaction Status-The Women's Health Initiative - Observational Study

	Satisfied sexually (N=35,719)	Not satisfied sexually (N=10,806)	
Sociodemographic variable	N (%)	N (%)	P-Value
Age (years)			
50–59	14,741 (75)	4,906 (25)	<.0001
60–69	15,673 (78)	4,513 (22)	
70–79	5,305 (79)	1,387 (21)	
Marital status			
Never married	345 (68)	161 (32)	<.0001
Divorced/separated	2,632 (67)	1,295 (33)	
Widowed	1,508 (72)	580 (28)	
Married/partnered	31,083 (78)	8,720 (22)	
Family income			
<$10,000	540 (70)	227 (30)	<.0001
$10,000–19,999	1,781 (73)	647 (37)	
$20,000–34,999	5,789 (75)	1,931 (25)	
$35,000–49,999	6,782 (77)	2,061 (23)	
$50,000–74,999	8,221 (77)	2,476 (23)	
$75,000+	10,191 (79)	2,782 (21)	
Don't know	895 (76)	281 (24)	
Race/ethnicity			
American Indian	142 (83)	29 (17)	<.0001
Asian/Pacific Islander	1,001 (85)	183 (16)	
Black/African American	2,269 (75)	776 (26)	
Hispanic	1,252 (77)	385 (24)	
White	30,629 (77)	9,325 (23)	
Unknown	347 (80)	87 (20)	
Education			
0–8 years	338 (77)	104 (24)	0.02
Some high school	887 (75)	291 (25)	
HS diploma/GED	5,476 (77)	1,603 (23)	
School after HS	12,658 (76)	4,002 (24)	
College degree or higher	16,092 (77)	4,726 (23)	
Employment status			
Technical/sales/admin	9,680 (76)	3,010 (24)	0.03
Service/labor	2,684 (77)	823 (24)	
Homemaker	6,934 (78)	1,953 (22)	
Managerial/professional	13,015 (77)	3,993 (24)	
Other	1,926 (76)	604 (24)	
Sexual orientation			
Heterosexual	34,771 (77)	10,497 (23)	0.10
Bisexual	248 (71)	103 (29)	
Lifetime lesbian	91 (77)	28 (24)	
Adult lesbian	116 (75)	39 (25)	
Never had sex	26 (72)	10 (28)	
Prefer not to respond	312 (79)	83 (21)	

We hypothesized that a model developed from these variables could accurately discriminate between the sexually satisfied and dissatisfied participants.

Statistical Analysis

Before creating a dichotomous outcome variable for sexual satisfaction, we verified that the extreme responses to the sexual satisfaction question (responses 1 and 4) reflected similar direction with greater magnitude when compared to their nearest respective middle range responses (responses 2 and 3). We then created a binary response variable (sexually satisfied versus unsatisfied.)

For bivariate analyses, we used t-tests for continuous variables and chi-squared tests for categorical variables, with alpha=0.05 to determine statistical significance. To assess the clinical significance of reported differences in continuous variables, we calculated an effect size, measured as the absolute value of the differences between the mean scores of those satisfied and not satisfied women divided by the standard deviation of the reference group (satisfied).[50] We use previously established guidelines for effect size[51] to describe clinical significance.

In order to test how robustly each group of variables independently discriminated between sexually satisfied and unsatisfied subjects, we created sequential logistic regression models with block entry for each category (model 1: sociodemographic, model 2: adds physical health, model 3: adds mental health, model 4: adds gynecologic variables, model 5: adds medications, and model 6: adds health behaviors). We report c-statistics for each model. Variables that were not significant in bivariate analysis were excluded from these models.

To test the robustness of our models, backwards and stepwise selection procedures were utilized, with both 0.05 and 0.20 selected as entry and retention criteria, respectively. The variables in the final model were the same for stepwise and backward selection, confirming the robustness of the results.

RESULTS

Fifty-two percent (48,300) of the respondents reported that they had been sexually active with another person in the past year. Of these, 96% (46,525) answered the sexual satisfaction question. Overall, 77% (35,719) reported satisfaction with sexual activity.

As shown in Table 1, sexual satisfaction was associated with increasing age and higher family income. Marital or partnered status was associated with sexual satisfaction, with more married or partnered participants reporting sexual satisfaction, versus all other participants. Identification with certain racial or ethnic groups was associated with sexual satisfaction. Other demographic variables showed statistically significant differences, but a small absolute difference. Sexual orientation was not associated with sexual satisfaction.

To ensure that the race and age findings were not an artifact of non-response bias, all non-responders were categorized as not satisfied. There was little change in the results among the groups identifying as American Indian or Asian. Similarly, when non-responders were categorized as not satisfied by age, the results were attenuated, but consistent in direction with the results reported in Table 1. Thus, differential non-response by either race or age did not account for our findings.

Table 2 shows the SF-36 subscales relating to physical and mental health. Among the physical health subscales, the effect sizes range from 0.09 to 0.14, a clinical difference that is insignificant. For the mental health subscales, the effect sizes range from 0.21 to 0.38. These differences are small but clinically meaningful.

Table 3 lists the gynecologic variables, medications, and health behaviors selected for investigation. Subjects with a history of oral contraceptive use were less likely to report sexual satisfaction, but this difference was small (76% versus 78%). Sexual satisfaction was equally proportional among participants reporting a history of gynecological cancer and those who did not. There was little difference in satisfaction among subjects reporting a history of hysterectomy (76%) compared to those who had never had a hysterectomy (77%). With respect to parity, our results were statistically significant, but we did not find a clear linear trend, threshold effect, or a j- or u-shaped relationship between parity and sexual satisfaction.

SSRI users were less likely to report sexual satisfaction than nonusers (66% versus 77%). Hormone therapy users and nonusers were equally likely to report sexual satisfaction. Satisfaction with sexual activity was associated with more exercise, never smoking, and normal body mass index (BMI). Alcohol users and nonusers reported equal rates of satisfaction.

Table 4 shows the results of sequential models with independent variables examined for association with sexual satisfaction. Model 1 includes demographics only and yields a c-statistic of 0.559, reflecting a limited ability to discriminate between satisfied and dissatisfied participants. In model 2, the

Table 2. Physical and Mental Health-Related Quality of Life by Sexual Satisfaction Status – The Women's Health Initiative – Observational Study

	Satisfied sexually (N=35,719)	Not satisfied sexually (N=10,806)	P-Value	Effect size
	Mean (SD*)	Mean (SD)		
Physical health subscales				
Physical functioning	85.02 (17.65)	83.47 (18.67)	<.0001	0.09
Role limitations due to physical health	77.54 (33.83)	73.54 (35.78)	<.0001	0.12
Bodily pain	76.63 (22.02)	73.77 (23.50)	<.0001	0.13
General health	76.32 (17.26)	73.95 (18.60)	<.0001	0.14
Mental health subscales				
Vitality	66.56 (17.98)	62.14 (20.06)	<.0001	0.25
Social functioning	91.66 (16.14)	88.19 (19.41)	<.0001	0.21
Role limitations due to emotional health	87.75 (26.40)	81.06 (31.78)	<.0001	0.25
Mental health	80.95 (12.87)	76.10 (15.91)	<.0001	0.38

SD = Standard deviation

Table 3. Gynecologic Variables, Medications and Health Behaviors by Sexual Satisfaction Status – The Women's Health Initiative – Observational Study

	Satisfied sexually (N=35,719)	Not satisfied sexually (N=10,806)	P-Value
	N (%)	N (%)	
Gynecological variables			
Oral contraceptive use ever			<.0001
Yes	16,964 (76)	5,447 (24)	
No	18,755 (78)	5,359 (22)	
History of gynecological cancer			0.11
Yes	1,021 (75)	341 (25)	
No	34,465 (77)	10,390 (23)	
History of hysterectomy			0.01
Yes	14,913 (76)	4,659 (24)	
No	20,768 (77)	6,135 (23)	
Parity			<.0001
Never pregnant	2,183 (77)	641 (23)	
Never had a term pregnancy	775 (72)	307 (28)	
1	2,935 (76)	928 (24)	
2	10,177 (77)	3,105 (23)	
3	9,451 (77)	2,801 (23)	
4	5,551 (78)	1,549 (22)	
5+	4,458 (76)	1,398 (24)	
Medications			
Selective serotonin reuptake inhibitor use			<.0001
Yes	1,111 (66)	578 (34)	
No	34,607 (77)	10,228 (23)	
Hormone therapy use			0.13
Yes	15,572 (76)	4,800 (24)	
No	20,146 (77)	6,006 (23)	
Health behaviors			
Exercise			<.0001
No activity	3,848 (74)	1,360 (26)	
Some activity	12,394 (77)	3,774 (23)	
2 to <4 episodes per week	6,974 (77)	2,079 (23)	
≥4 episodes per week	12,193 (78)	3,516 (22)	
Smoking			<.0001
Never smoked	18,897 (78)	5,208 (22)	
Past smoker	14,616 (75)	4,813 (25)	
Current smoker	1,825 (74)	654 (26)	
Body mass index			<.001
<18.5 Underweight	349 (74)	122 (26)	
18.5–24.9 Normal	15,583 (77)	4,565 (23)	
25.0–24.9 Overweight	12,096 (77)	3,657 (23)	
≥30 Obesity	7,301 (76)	2,343 (24)	
Alcohol use			0.54
Yes	25,341 (76)	7,880 (24)	
No	6,246 (77)	1,908 (23)	

general health and pain constructs of the SF-36 met our model retention criteria. With addition of the mental health subscales in model 3, the physical health subscales were excluded, and the mental health and vitality constructs were retained. In model 4, adding gynecologic variables, parity and oral contraceptive use were retained. In model 5, SSRI use was retained. In model 6, smoking status was retained. With each sequential model, the c-statistic improves only modestly, to a final c-statistic of 0.613, reflecting a limited ability to discriminate between women who report sexual satisfaction and those that do not.

In model 6, among the demographic variables, the strongest association with sexual satisfaction was found with race or ethnicity. American Indians or Alaskan Natives were 2.7 times more likely to be sexually satisfied than whites, followed by Asian (OR=1.5) and Hispanic (OR=1.1) respondents. Being married or partnered was significantly associated with sexual satisfaction, compared to all other participants. The oldest age cohort, 70–79, had greater odds of satisfaction, compared to the younger cohorts.

Among the SF-36 subscales, two scales associated with mental health were retained, with greater odds of sexual satisfaction associated with greater mental health (OR=1.3) and greater vitality (OR=1.1). Certain parous states were associated with sexual satisfaction. No prior use of oral contraceptives was associated with greater odds of sexual satisfaction (OR=1.1), as was no SSRI use (OR=1.4). Former smokers were less likely to report satisfaction (OR=0.8) versus never smokers.

Despite excluding subjects who had not had sex with a partner in the past year, satisfaction with sexual frequency could confound the overall sexual satisfaction construct. In bivariate analysis, satisfaction with sexual frequency was highly associated ($p<.0001$) with satisfaction with sexual activity. Participants who would like sex less often, more often, and who preferred not to answer were more likely to be sexually dissatisfied than participants who were satisfied with current sexual frequency.

To test whether satisfaction with sexual frequency confounded our overall results with respect to sexual satisfaction, we reran model 6 including satisfaction with sexual frequency. All variables in the model were the same as the variables in Table 3, except for age and oral contraceptive use, which no longer met our retention criterion, and satisfaction with sexual frequency, which met our retention criterion and was therefore included.

Because this could indicate that age may function as a surrogate for satisfaction with sexual frequency, we assessed the bivariate association between age and satisfaction with sexual frequency. Among dissatisfied participants, a higher proportion of women aged 50–59 would like sex more often than women aged 70–79 (35% versus 20%), and a higher proportion of women aged 70–79 would like sex less often than women aged 50–59 (6% versus 5%).

DISCUSSION

Postmenopausal women in the Women's Health Initiative Observation Cohort reported sexual satisfaction if they were older, married or partnered, emotionally healthier, not taking SSRIs, and never smokers. Sexual satisfaction was also associated with identification with certain racial or ethnic groups and showed a complex relationship with parity. While many of these factors are not modifiable, several, including mental health status and SSRI use, are amenable to primary care intervention. After inclusion of all of the variables, among postmenopausal women who had sex with a partner in the past year, our final model had limited predictive ability. This reflects the complexity of the satisfaction construct, residual confounding, and unmeasured variables.

Our age findings were unexpected. Previous studies describe increasing incidence of dyspareunia with aging, thought to be a strong correlate of sexual dissatisfaction.[5,6,52] However,

Table 4. Odds of Satisfaction with Sexual Activity Among Sexually Active Postmenopausal Women–The Women's Health Initiative–observational study

Model	Variable	Estimate (95% CI)	P-Value	C-Statistic	Change in c-statistic
Model 1: Demographics only	Age		<.0001	0.559	–
	50–59	0.723 (0.649, 0.805)			
	60–69	0.844 (0.761, 0.936)			
	70–79	1.00			
	Race/ethnicity		0.0035		
	American Indian or Alaskan Native	2.105 (1.037, 4.272)			
	Asian or Pacific Islander	1.707 (1.266, 2.300)			
	Black or African-American	1.037 (0.893, 1.203)			
	Hispanic/Latino	1.079 (0.882, 1.319)			
	Other	1.170 (0.810, 1.690)			
	White (not of Hispanic origin)	1.00			
	Marital status		<.0001		
	Divorced or separated	0.614 (0.546, 0.691)			
	Never married	0.521 (0.388, 0.700)			
	Widowed	0.739 (0.630, 0.867)			
	Married/partnered	1.00			
	Family income		0.0405		
	$10,000–19,999	0.953 (0.711, 1.276)			
	$20,000–34,999	0.956 (0.727, 1.256)			
	$35,000–49,999	1.086 (0.825, 1.428)			
	$50,000–74,999	1.033 (0.786, 1.358)			
	$75,000+	1.113 (0.845, 1.465)			
	Don't know	0.888 (0.637, 1.239)			
	<$10,0000	1.00			
Model 2: Demographics and physical health	Age		<.0001	0.574	0.015
	50–59	0.732 (0.658, 0.813)			
	60–69	0.842 (0.759, 0.934)			
	70–79	1.00			
	Race/ethnicity		0.0029		
	American Indian or Alaskan Native	2.329 (1.145, 4.739)			
	Asian or Pacific Islander	1.640 (1.217, 2.211)			
	Black or African-American	1.083 (0.932, 1.260)			
	Hispanic/Latino	1.099 (0.898, 1.345)			
	Other	1.185 (0.816, 1.719)			
	White (not of Hispanic origin)	1.00			
	Marital status		<.0001		
	Divorced or separated	0.587 (0.524, 0.657)			
	Never married	0.487 (0.363, 0.655)			
	Widowed	0.709 (0.606, 0.830)			
	Married/partnered	1.00			
	General health	1.114 (1.073, 1.156)	<.0001		
	Bodily pain	1.076 (1.037, 1.117)	0.0001		
Model 3: Demographics, physical health and mental health	Age		0.0004	0.605	0.031
	50–59	0.805 (0.723, 0.896)			
	60–69	0.863 (0.777, 0.959)			
	70–79	1.00			
	Race/ethnicity		0.0020		
	American Indian or Alaskan Native	2.711 (1.279, 5.745)			
	Asian or Pacific Islander	1.607 (1.189, 2.171)			
	Black or African-American	1.034 (0.888, 1.203)			
	Hispanic/Latino	1.150 (0.937, 1.411)			
	Other	1.236 (0.845, 1.808)			
	White (not of Hispanic origin)	1.00			
	Marital status		<.0001		
	Divorced or separated	0.621 (0.554, 0.696)			
	Never married	0.514 (0.381, 0.692)			
	Widowed	0.764 (0.651, 0.896)			
	Married/partnered	1.00			
	Mental health	1.304 (1.255, 1.355)	<.0001		
	Vitality	1.083 (1.042, 1.127)	<.0001		

(*continued on next page*)

Table 4. (continued)

Model	Variable	Estimate (95% CI)	P-Value	C-Statistic	Change in c-statistic
Model 4: Demographics, physical health, mental health and gynecologic variables	Age		0.0098	0.608	0.003
	50–59	0.839 (0.749, 0.940)			
	60–69	0.883 (0.794, 0.982)			
	70–79	1.00			
	Race/ethnicity		0.0025		
	American Indian or Alaskan Native	2.705 (1.275, 5.740)			
	Asian or Pacific Islander	1.596 (1.181, 2.157)			
	Black or African-American	1.027 (0.881, 1.196)			
	Hispanic/Latino	1.150 (0.937, 1.412)			
	Other	1.222 (0.835, 1.787)			
	White (not of Hispanic origin)	1.00			
	Marital status		<.0001		
	Divorced or separated	0.623 (0.555, 0.698)			
	Never married	0.477 (0.347, 0.655)			
	Widowed	0.759 (0.647, 0.890)			
	Married/partnered	1.00			
	Mental health	1.303 (1.254, 1.354)	<.0001		
	Vitality	1.083 (1.041, 1.126)	<.0001		
	Parity		0.0431		
	1	0.994 (0.823, 1.199)			
	2	0.860 (0.734, 1.008)			
	3	0.854 (0.728, 1.002)			
	4	0.920 (0.776, 1.091)			
	5+	0.824 (0.692, 0.980)			
	Never had term pregnancy	0.763 (0.593, 0.983)			
	Never pregnant	1.00			
	Oral contraceptive use ever		0.0044		
	No	1.109 (1.033, 1.191)			
	Yes	1.00			
Model 5: Demographics, physical health, mental health, gynecologic variables, and medications	Age		0.0136	0.610	0.002
	50–59	0.844 (0.754, 0.945)			
	60–69	0.885 (0.796, 0.984)			
	70–79	1.00			
	Race/ethnicity		0.0032		
	American Indian or Alaskan Native	2.710 (1.277, 5.749)			
	Asian or Pacific Islander	1.580 (1.169, 2.135)			
	Black or African-American	1.019 (0.875, 1.187)			
	Hispanic/Latino	1.143 (0.931, 1.403)			
	Other	1.214 (0.830, 1.775)			
	White (not of Hispanic origin)	1.00			
	Marital status		<.0001		
	Divorced or separated	0.624 (0.556, 0.700)			
	Never married	0.478 (0.348, 0.656)			
	Widowed	0.759 (0.647, 0.891)			
	Married/partnered	1.00			
	Mental health	1.297 (1.249, 1.348)	<.0001		
	Vitality	1.074 (1.033, 1.117)	0.0004		
	Parity		0.0442		
	1	0.987 (0.818, 1.191)			
	2	0.857 (0.731, 1.004)			
	3	0.851 (0.725, 0.998)			
	4	0.915 (0.772, 1.086)			
	5+	0.818 (0.687, 0.973)			
	Never had term pregnancy	0.766 (0.594, 0.987)			
	Never pregnant	1.00			
	Oral contraceptive use ever		0.0054		
	No	1.106 (1.030, 1.188)			
	Yes	1.00			
	Selective serotonin reuptake inhibitor use		<.0001		
	No	1.381 (1.177, 1.620)			
	Yes	1.00			

(continued on next page)

Table 4. (continued)

Model	Variable	Estimate (95% CI)	P-Value	C-Statistic	Change in c-statistic
Model 6: Demographics, physical health, mental health, gynecologic variables, medications, behaviors	Age		0.0202	0.613	0.003
	50–59	0.851 (0.759, 0.953)			
	60–69	0.892 (0.802, 0.992)			
	70–79	1.00			
	Race/ethnicity		0.0068		
	American Indian or Alaskan Native	2.678 (1.262, 5.681)			
	Asian or Pacific Islander	1.524 (1.128, 2.061)			
	Black or African-American	1.011 (0.867, 1.178)			
	Hispanic/Latino	1.123 (0.915, 1.380)			
	Other	1.220 (0.834, 1.785)			
	White (not of Hispanic origin)	1.00			
	Marital status		<.0001		
	Divorced or separated	0.631 (0.563, 0.708)			
	Never married	0.485 (0.353, 0.666)			
	Widowed	0.766 (0.653, 0.899)			
	Married/partnered	1.00			
	Mental health	1.292 (1.243, 1.343)	<.0001		
	Vitality	1.079 (1.038, 1.123)	0.0002		
	Parity		0.0402		
	1	0.988 (0.819, 1.192)			
	2	0.854 (0.729, 1.002)			
	3	0.848 (0.723, 0.996)			
	4	0.910 (0.767, 1.080)			
	5+	0.811 (0.681, 0.966)			
	Never had term pregnancy	0.775 (0.601, 0.998)			
	Never pregnant	1.00			
	Oral contraceptive use ever		0.0107		
	No	1.097 (1.022, 1.179)			
	Yes	1.00			
	Selective serotonin reuptake inhibitor use		0.0001		
	No	1.369 (1.167, 1.607)			
	Yes	1.00			
	Smoking		<.0001		
	Current smoker	0.905 (0.778, 1.053)			
	Past smoker	0.834 (0.778, 0.894)			
	Never smoked	1.00			

many of these studies focused on premenopausal women, or compared premenopausal to postmenopausal women. Our cohort of postmenopausal women was likely more homogeneous with respect to hormonal milieu. Further research, including research on lubricant use among different ages of postmenopausal women, may better elaborate this association.

Additionally, sexually-related distress,[14] and anxiety about sexual performance[6] and attractiveness decrease with age.[53] As satisfaction may be impeded by anxiety over sexual performance or desirability, relative satisfaction may increase with age.

Selection bias may partially explain satisfaction with increasing age. All of the women in this analysis were healthy enough to have sexual activity with a partner in the past year. With increasing age, the presence of a sexual partner may be increasingly protective against dissatisfaction. Stability of partnered relationships is protective against sexual dysfunction.[6]

Decreased satisfaction in women with five or more births may reflect the late urogenital sequelae of pregnancy and delivery. Pelvic organ prolapse and urinary incontinence may be associated with increased parity,[54–56] and may result in decreased sexual satisfaction. However, those who had conceived but never carried to term were least likely to report sexual satisfaction. Although this may reflect long-standing sequelae of pelvic disease or endocrinopathies that can interfere with gestation, it is unclear that this would remain associated with sexual satisfaction after menopause. This finding deserves further investigation.

None of the SF-36 subscales that aggregate most strongly with physical health were retained in our final model. This result is disconcordant with other studies.[6,57] Of note, the vitality measure retained in our final model cross-correlates with the physical health aggregate. Additionally, health and sexuality are more strongly linked for men versus women.[29]

Mental health was a strong predictor of lower satisfaction with sexual activity, in this study and others.[6,57] Because the WHI-OS excluded women with major depression or mental health disorders, enrolling overall healthy women,[18] the range of mental health symptoms many be mild compared to population-based data. Controlling for mental health status, SSRI use remained a strong correlate of decreased sexual satisfaction. Women may have more severe SSRI sexual side effects, including orgasm delay, than men.[41] In preliminary work, we controlled specifically for depressive symptoms and found a similar relationship with SSRI use.

Our findings with respect to use of oral contraceptives were surprising. Detrimental effects of oral contraceptives on sexual functioning may be sustained after discontinuation of the medications.[37] However, removal of this variable from the final model with addition of the sexual frequency construct suggests this finding lacks robustness.

There was no difference between current and never smokers in sexual satisfaction. Among former smokers we found a modest but statistically significant decrement in sexual satisfaction. We explored whether this finding reflected a "healthy smoker" effect, in which individuals who take up and persist in smoking are relatively resistant to the adverse health effects of smoking.[58] The mean overall health scores of former smokers (76.3) was higher than that of current smokers (72.5, p<.0001). Thus, this finding may reflect type I error.

Despite previous randomized controlled trials showing estrogen therapy effects on sexual satisfaction,[59] hormone therapy did not have a significant impact on sexual satisfaction. Confounding by indication suggests that women with sexual dissatisfaction due to atrophic vaginitis may take exogenous estrogen. Conversely, women not on exogenous estrogen are less likely to have atrophic vaginitis.

Our study included only those women who reported partnered sexual activity. This sampling scheme excluded women who may have been sexually active, but were not sexually active with another person. Although sexual satisfaction is not dependent upon the presence of a partner, this exclusion was necessary to decrease confounding of our satisfaction construct by lack of a sexual partner.[6] Additionally, it isolated those who were sexually active and satisfied versus those who were not sexually active, but remained satisfied. The relative importance of sex decreases among older women,[14] and sexual interest and desire decrease after menopause.[9,60,61] Thus, respondents who had no sexual activity may have been very satisfied. Examination of sexual satisfaction among unpartnered women was beyond the scope of this study.

This study has several strengths. The Women's Health Initiative is a large cohort and has good representation of women across geographic, racial, and ethnic groups. This is the first study to describe the correlates of sexual satisfaction among postmenopausal women and to examine the individual subscales of the SF-36 in relation to sexual satisfaction.

An important limitation of this work is that the sexual satisfaction construct has not been validated. Of note, the question has strong face validity and similar wording to questions on validated instruments. Additionally, longitudinal data on sexual satisfaction were not available. Thus, changes in associations over time could not be described, nor could baseline factors that might predict change in sexual satisfaction. Further, we cannot exclude the possibility that secular trends in the correlates of sexual satisfaction exist, and that younger cohorts may report a different set of correlates once they reach menopause.

Despite our exclusion of those without sex with a partner in the last year, subjects could answer the satisfaction question based on quantity of sex. A sensitivity analysis that included satisfaction with sexual frequency found that satisfaction with sexual frequency replaced age in multivariate analysis, suggesting the centrality of sexual frequency to sexual satisfaction. Further research should better correlate sexual frequency with sexual satisfaction among postmenopausal women.

CONCLUSIONS

Satisfaction with sexual activity among postmenopausal women was associated with demographic and historical factors that are not amenable to physician intervention. However, it is also associated with potentially modifiable factors, including self-reported mental health status and SSRI use. These results should be interpreted with caution, as the final model had limited ability to discriminate between the satisfied and dissatisfied participants. Further research may better elaborate the cofactors associated with sexual satisfaction among postmenopausal women.

Acknowledgements: Funding/Support: *The Women's Health Initiative program was funded by the National Heart, Lung, and Blood Institute of the National Institutes of Health, US Department of Health and Human Services. Dr. McCall-Hosenfeld was supported by a Department of Veterans Affairs Special Fellowship in the Health Issues of Women Veterans when this research was performed.*

Role of the Sponsor: *The funding organization had representation on the steering committee, which governed the design and conduct of the study, the interpretation of the data, and the preparation and approval of manuscript. The National Heart, Lung, and Blood Institute Program Office reviewed the manuscript prior to publication.*

WHI Investigators by Clinical Center: Program Office*: (National Heart, Lung, and Blood Institute, Bethesda, MD) Elizabeth Nabel, Jacques Rossouw, Shari Ludlam, Linda Pottern, Joan McGowan, Leslie Ford, and Nancy Geller.*

Clinical Coordinating Center*: (Fred Hutchinson Cancer Research Center, Seattle, WA)*

Ross Prentice, Garnet Anderson, Andrea LaCroix, Charles L. Kooperberg, Ruth E. Patterson, Anne McTiernan; (Wake Forest University School of Medicine, Winston-Salem, NC) Sally Shumaker; (Medical Research Labs, Highland Heights, KY) Evan Stein; (University of California at San Francisco, San Francisco, CA) Steven Cummings.

Clinical Centers*: (Albert Einstein College of Medicine, Bronx, NY) Sylvia Wassertheil-Smoller; (Baylor College of Medicine, Houston, TX) Aleksandar Rajkovic; (Brigham and Women's Hospital, Harvard Medical School, Boston, MA) JoAnn Manson; (Brown University, Providence, RI) Annlouise R. Assaf; (Emory University, Atlanta, GA) Lawrence Phillips; (Fred Hutchinson Cancer Research Center, Seattle, WA) Shirley Beresford; (George Washington University Medical Center, Washington, DC) Judith Hsia; (Los Angeles Biomedical Research Institute at Harbor- UCLA Medical Center, Torrance, CA) Rowan Chlebowski; (Kaiser Permanente Center for Health Research, Portland, OR) Evelyn Whitlock; (Kaiser Permanente Division of Research, Oakland, CA) Bette Caan; (Medical College of Wisconsin, Milwaukee, WI) Jane Morley Kotchen; (MedStar Research Institute/Howard University, Washington, DC) Barbara V. Howard; (Northwestern University, Chicago/Evanston, IL) Linda Van Horn; (Rush Medical Center, Chicago, IL) Henry Black; (Stanford Prevention Research Center, Stanford, CA) Marcia L. Stefanick; (State University of New York at Stony Brook, Stony Brook, NY) Dorothy Lane; (The Ohio State University, Columbus, OH) Rebecca Jackson; (University of Alabama at Birmingham, Birmingham, AL) Cora E. Lewis; (University of Arizona, Tucson/Phoenix, AZ) Tamsen Bassford; (University at Buffalo, Buffalo, NY) Jean Wactawski-Wende; (University of California at Davis, Sacramento, CA) John Robbins; (University of California at Irvine, CA) F. Allan Hubbell; (University of California at Los Angeles, Los Angeles, CA) Lauren Nathan; (University of California at San Diego, LaJolla/Chula Vista, CA) Robert D. Langer; (University of Cincinnati, Cincinnati, OH) Margery Gass; (University of Florida, Gainesville/Jacksonville, FL) Marian Limacher; (University of Hawaii, Honolulu, HI) David Curb; (University of Iowa, Iowa City/Davenport, IA) Robert Wallace; (University of Massachusetts/Fallon Clinic, Worcester, MA) Judith Ockene; (University of Medicine and Dentistry of New Jersey, Newark, NJ) Norman Lasser; (University of Miami, Miami, FL) Mary Jo O'Sullivan; (University of Minnesota, Minneapolis, MN) Karen Margolis; (University of Nevada, Reno, NV) Robert Brunner; (University of North Carolina, Chapel Hill, NC) Gerardo Heiss; (University of Pittsburgh, Pittsburgh, PA) Lewis Kuller; (University of Tennessee, Memphis, TN) Karen C. Johnson; (University of Texas Health Science*

Center, San Antonio, TX) Robert Brzyski; (University of Wisconsin, Madison, WI) Gloria E. Sarto; (Wake Forest University School of Medicine, Winston-Salem, NC) Mara Vitolins; (Wayne State University School of Medicine/Hutzel Hospital, Detroit, MI) Susan Hendrix.

Conflict of Interest: *None disclosed.*

Corresponding Author: *Jennifer S. McCall-Hosenfeld, MD, MSc; Division of General Internal Medicine and Department of Public Health Sciences, Pennsylvania State University College of Medicine, Hershey, PA, USA (e-mail: jsm31@psu.edu).*

REFERENCES

1. **Rosen, R, Brown C, Heiman J, et al.** The Female Sexual Function Index (FSFI): a multidimensional self–report instrument for the assessment of female sexual function. J Sex Marital Ther. 2000;26(2):191–208.
2. **Rosen, RC, Lobo RA, Block BA, Yang H-M, Zipfel LM.** Menopausal Sexual Interest Questionnaire (MSIQ): a unidimensional scale for the assessment of sexual interest in postmenopausal women. J Sex Marital Ther. 2004;30(4):235–50.
3. **Rust, J, Golombok S.** The GRISS: A Psychometric Instrument for the Assessment of Sexual Dysfunction. Arch Sex Behav. 1986;15(2):157–65.
4. **Nicolosi, A, Laumann EO, Glasser DB, et al.** Sexual behavior and sexual dysfunctions after age 40: the global study of sexual attitudes and behaviors. Urology. 2004;64(5):991–7.
5. **Laumann, EO, Nicolosi A, Glasser DB, et al.** Sexual problems among women and men aged 40–80 y: prevalence and correlates identified in the Global Study of Sexual Attitudes and Behaviors. Int J Impot Res. 2005;17(1):39–57.
6. **Laumann, EO, Paik A, Rosen RC.** Sexual dysfunction in the United States: prevalence and predictors. JAMA. 1999;281(6):537–44.
7. **Salonia, A, Munarriz RM, Naspro R, et al.** Women's sexual dysfunction: a pathophysiological review. BJU Int. 2004;93(8):1156–64.
8. **Nazareth, I, Boynton P, King M.** Problems with sexual function in people attending London general practitioners: cross sectional study. BMJ. 2003;327(7412):423.
9. **Dennerstein, L, Smith A, Morse C.** Sexuality and the menopause. J Psychosomatic Obstet Gynecol. 1994;15:59–66.
10. **Maas, CP, Weijenborg PTM, ter Kuile MM.** The effect of hysterectomy on sexual functioning. Annu Rev Sex Res. 2003;14:83–113.
11. **Dennerstein, L, Koochaki P, Barton I, Graziottin A.** Hypoactive sexual desire disorder in menopausal women: a survey of Western European women. J Sex Med. 2006;3(2):212–22.
12. **Dennerstein, L, Lehert P.** Modeling mid-aged women's sexual functioning: a prospective, population-based study. J Sex Marital Ther. 2004;30 (3):173–83.
13. **Dennerstein, L, Hayes RD.** Confronting the challenges: epidemiological study of female sexual dysfunction and the menopause. J Sex Med. 2005;2(Suppl 3):118–32.
14. **Hayes, R, Dennerstein L.** The impact of aging on sexual function and sexual dysfunction in women: a review of population-based studies. J Sex Med. 2005;2(3):317–30.
15. **Cain, VS, Johannes CB, Avis NE, et al.** Sexual functioning and practices in a multi-ethnic study of midlife women: baseline results from SWAN. J Sex Res. 2003;40(3):266–76.
16. **Howard, BV, Kuller L, Langer R, et al.** Risk of cardiovascular disease by hysterectomy status, with and without oophorectomy: the Women's Health Initiative Observational Study. Circulation. 2005;111(12):1462–70.
17. **Hsia, J, Aragaki A, Bloch M, LaCroix AZ, Wallace R, Investigators WHI.** Predictors of angina pectoris versus myocardial infarction from the Women's Health Initiative Observational Study. Am J Cardiol. 2004;93 (6):673–8.
18. **Langer, RD, White E, Lewis CE, Kotchen JM, Hendrix SL, Trevisan M.** The Women's Health Initiative Observational Study: baseline characteristics of participants and reliability of baseline measures. Ann Epidemiol. 2003;13(9 Suppl):S107–21.
19. **Hays, J, Hunt JR, Hubbell FA, et al.** The Women's Health Initiative recruitment methods and results. Ann Epidemiol. 2003;13(9 Suppl):S18–77.
20. **Stefanick, ML, Cochrane BB, Hsia J, Barad DH, Liu JH, Johnson SR.** The Women's Health Initiative postmenopausal hormone trials: overview and baseline characteristics of participants. Ann Epidemiol. 2003;13(9 Suppl):S78–86.
21. **Valanis, BG, Bowen DJ, Bassford T, Whitlock E, Charney P, Carter RA.** Sexual orientation and health: comparisons in the women's health initiative sample. Arch Fam Med. 2000;9(9):843–53.
22. **Brunner RL, Gass M, Aragaki A, et al.** Effects of conjugated equine estrogen on health-related quality of life in postmenopausal women with hysterectomy: results from the Women's Health Initiative Randomized Clinical Trial. Arch Intern Med. 2005:1976–86.
23. **Hays, J, Ockene JK, Brunner RL, et al.** Effects of estrogen plus progestin on health-related quality of life. N Engl J Med. 2003;348(19):1839–54.
24. **Oksuz, E, Malhan S, Oksuz E, Malhan S.** Prevalence and risk factors for female sexual dysfunction in Turkish women. J Urol. 2006;175(2):654–8.
25. **Lewis, RW, Fugl-Meyer KS, Bosch R, et al.** Epidemiology/risk factors of sexual dysfunction. J Sex Med. 2004;1(1):35–9.
26. **Addis, IB, Ireland CC, Vittinghoff E, Lin F, Stuenkel CA, Hulley S.** Sexual activity and function in postmenopausal women with heart disease. Obstet Gynecol. 2005;106(1):121–7.
27. **Henderson AW, Lehavot K, Simoni JM.** Ecological Models of Sexual Satisfaction among Lesbian/Bisexual and Heterosexual Women. Arch Sex Behav 2008.
28. **Tracy, JK, Junginger J.** Correlates of lesbian sexual functioning. J Women's Health. 2007;16(4):499–509.
29. **Lindau, ST, Schumm LP, Laumann EO, Levinson W, O'Muircheartaigh CA, Waite LJ.** A study of sexuality and health among older adults in the United States. N Engl J Med. 2007;357(8):762–74.
30. **Basson, R, Schultz WW.** Sexual sequelae of general medical disorders. Lancet. 2007;369(9559):409–24.
31. **Nappi, R, Salonia A, Traish AM, et al.** Clinical biologic pathophysiologies of women's sexual dysfunction. J Sex Med. 2005;2(1):4–25.
32. **Ware, J, Sherbourne C.** The MOS 36-Item Short-Form Health Survey (SF-36), I: conceptual framework and item selection. Medical Care. 1992;30:473–83.
33. **Hays, R, Sherbourne C, Mazel R.** The RAND 36-Item Health Survey 1.0. Health Econ. 1993;2(3):213–5.
34. **Bartlik, B, Kocsis JH, Legere R, et al.** Sexual dysfunction secondary to depressive disorders. J Gend Specif Med. 1999;2(2):52–60.
35. **Kennedy, SH, Dickens SE, Eisfeld BS, et al.** Sexual dysfunction before antidepressant therapy in major depression. J Affect Disord. 1999;56(2–3):201–8.
36. **Dobkin, RD, Leiblum SR, Rosen RC, et al.** Depression and sexual functioning in minority women: current status and future directions. J Sex Marital Ther. 2006;32(1):23–36.
37. **Panzer, C, Wise S, Fantini G, et al.** Impact of oral contraceptives on sex hormone-binding globulin and androgen levels: a retrospective study in women with sexual dysfunction. J Sex Med. 2006;3(1):104–13.
38. **Andersen, BL, Woods XA, Copeland LJ.** Sexual self-schema and sexual morbidity among gynecologic cancer survivors. J Consult Clin Psychol. 1997;65(2):221–9.
39. **Pieterse, QD, Maas CP, ter Kuile MM, et al.** An observational longitudinal study to evaluate miction, defecation, and sexual function after radical hysterectomy with pelvic lymphadenectomy for early-stage cervical cancer. Int J Gynecol Cancer. 2006;16(3):1119–29.
40. **Botros, SM, Abramov Y, Miller J-JR, et al.** Effect of parity on sexual function: an identical twin study. Obstet Gynecol. 2006;107(4):765–70.
41. **Hensley, PL, Nurnberg HG.** SSRI sexual dysfunction: a female perspective. J Sex Marital Ther. 2002;28(Suppl 1):143–53.
42. **Rosen, RC, Lane RM, Menza M.** Effects of SSRIs on sexual function: a critical review. J Clin Psychopharmacol. 1999;19(1):67–85.
43. **Wassertheil-Smoller, S, Shumaker S, Ockene J, et al.** Depression and cardiovascular sequelae in postmenopausal women. The Women's Health Initiative (WHI). Arch Intern Med. 2004;164(3):289–98.
44. **Anderson, GL, Manson J, Wallace R, et al.** Implementation of the Women's Health Initiative study design. Ann Epidemiol. 2003;13(9 Suppl):S5–17.
45. **Davison, SL, Bell RJ, LaChina M, et al.** Sexual function in well women: stratification by sexual satisfaction, hormone use, and menopause status. J Sex Med. 2008;5(5):1214–22.
46. **Meston, CM, Gorzalka BB, Meston CM, Gorzalka BB.** Differential effects of sympathetic activation on sexual arousal in sexually dysfunctional and functional women. J Abnorm Psychol. 1996;105 (4):582–91.
47. **Esposito, K, Ciotola M, Giugliano F, et al.** Association of body weight with sexual function in women. Int J Impot Res. 2007;19(4):353–7.
48. **Covington, SS, Kohen J, Covington SS, Kohen J.** Women, alcohol, and sexuality. Adv Alcohol Subst Abuse. 1984;4(1):41–56.
49. **Peterson, JS, Hartsock N, Lawson G, Peterson JS, Hartsock N, Lawson G.** Sexual dissatisfaction of female alcoholics. Psychol Rep. 1984;55(3):744–6.

50. **Rosenthal R.** Parametric measures of effect size. In: Cooper H, Hedges LV, eds. The handbook of research synthesis. New York: Russell Sage Foundation; 1994:231–44.
51. **Cohen, J.** Statistical power analysis for the behavioral sciences. 2Hillsdale, NJ: L. Erlbaum; 1988.
52. **Kadri, N, McHichi Alami KH, McHakra Tahiri S.** Sexual dysfunction in women: population based epidemiological study. Arch Womens Ment Health. 2002;5(2):59–63.
53. **Richters, J, Grulich AE, de Visser RO, Smith AMA, Rissel CE.** Sex in Australia: autoerotic, esoteric and other sexual practices engaged in by a representative sample of adults. Aust N Z J Public Health. 2003;27 (2):180–90.
54. **Ghetti, C, Gregory WT, Clark AL.** Risk factors for surgically managed pelvic organ prolapse and urinary incontinence. Int J Gynaecol Obstet. 2007;98(1):63–4.
55. **Connolly, TJ, Litman HJ, Tennstedt SL, Link CL, McKinlay JB.** The effect of mode of delivery, parity, and birth weight on risk of urinary incontinence. Int Urogynecol J. 2007;18(9):1033–42.
56. **Hendrix, SL, Clark A, Nygaard I, Aragaki A, Barnabei V, McTiernan A.** Pelvic organ prolapse in the Women's Health Initiative: gravity and gravidity. Am J Obstet Gynecol. 2002;186(6):1160–6.
57. **Heiman, JR.** Sexual dysfunction: overview of prevalence, etiological factors, and treatments. J of Sex Res. 2002;39(1):73–8.
58. **Becklake, M, Lalloo U.** The 'healthy smoker': a phenomenon of health selection? Respiration. 1990;57(3):137–44.
59. **Alexander, JL, Kotz K, Dennerstein L, Kutner SJ, Wallen K, Notelovitz M.** The effects of postmenopausal hormone therapies on female sexual functioning: a review of double-blind, randomized controlled trials. Menopause. 2004;11(6 Pt 2):749–65.
60. **Avis, NE, Stellato R, Crawford S, Johannes C, Longcope C.** Is there an association between menopause status and sexual functioning? Menopause. 2000;7(5):297–309.
61. **Avis, NE, Zhao X, Johannes CB, Ory M, Brockwell S, Greendale GA.** Correlates of sexual function among multi-ethnic middle-aged women: results from the Study of Women's Health Across the Nation (SWAN). Menopause. 2005;12(4):385–98.

Published in final edited form as:
J Gerontol A Biol Sci Med Sci. 2008 September ; 63(9): 951–959.

Advance Care Planning and Health Care Preferences of Community-Dwelling Elders: The Framingham Heart Study

Ellen P. McCarthy, PhD, MPH[1], **Michael J. Pencina, PhD**[2,3], **Margaret Kelly-Hayes, EdD, RN**[2,4], **Jane C. Evans, DSc, MPH**[2,3], **Elizabeth J. Oberacker, BA**[2], **Ralph B. D'Agostino, PhD**[2,3], **Risa B. Burns, MD, MPH**[1], and **Joanne M. Murabito, MD, ScM**[2,5]

1*Division of General Medicine and Primary Care, Department of Medicine, Harvard Medical School, Beth Israel Deaconess Medical Center, Boston, Massachusetts.*

2*National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, Massachusetts.*

3*Department of Mathematics and Statistics, Statistical Consulting Unit, Boston University, Boston, Massachusetts.*

4*Department of Neurology, Boston University School of Medicine, Boston Medical Center, Boston, Massachusetts.*

5*Section of General Internal Medicine, Department of Medicine, Boston University School of Medicine, Boston Medical Center, Boston, Massachusetts.*

Abstract

Objectives—To describe self-reported advance care planning, health care preferences, use of advance directives, and health perceptions in a very elderly community-dwelling sample.

Methods—Surviving participants of the original cohort of the Framingham Heart Study who were cognitively intact and attended a routine research exam between February 2004 and October 2005. Participants were queried about discussions about end of life care, preferences for care, documentation of advance directives, and health perceptions.

Results—Among 220 community-dwelling respondents, 67% were women with a mean age of 88 years (range 84-100). Overall 69% discussed their wishes for medical care at the end of life with someone, but only 17% discussed their wishes with a physician or health care provider. Two-thirds had a health care proxy, 55% had a living will, and 41% had both. Most (80%) respondents preferred comfort care over life-extending care, and 71% preferred to die at home; however, substantially fewer respondents said they would rather die than receive specific life-prolonging interventions [chronic ventilator (63%) or feeding tube (64%)]. Many were willing to endure distressing health states, with less than half indicating that they would rather die than live out their life in a great deal of pain (46%) or be confused/forgetful (45%) all of the time.

Conclusions—Although the vast majority of very elderly community-dwellers in this sample appear to prefer comfort measures at the end of life, many said they were willing to endure specific life-prolonging interventions and distressing health states to avoid death. Our results highlight the need for physicians better understand patients' preferences and goals of care to help them make informed decisions at the end of life.

Corresponding Author and Requests for Reprints: Ellen P. McCarthy, PhD, Division of General Medicine and Primary Care, Department of Medicine, Beth Israel Deaconess Medical Center, 1309 Beacon Street, Suite 220, Brookline, MA 02446. Telephone Number (617) 754-1403; Fax Number (617) 754-1440. E-mail: E-mail: emccarth@bidmc.harvard.edu.

Conflict of Interest: The authors report no conflicts of interest.

This paper was presented in part at the annual meeting of the Society of General Internal Medicine, New Orleans, Lousiana, May 2005.

Keywords

advance directives; geriatrics; end of life care; patient centered care; decision-making

INTRODUCTION

Advance care planning is one key element to achieving patient autonomy by allowing patients' to participate in decisions about their medical care (1,2). However, many patients near the end of life lose decision-making capacity or become too ill to participate in such decisions (1). When this happens, patients' loved ones and physicians must make these difficult decisions, even though they may be unaware of or disagree with patient's preferences for care (3,4). Advance care planning is based on the premise that ongoing discussions about end of life issues accompanied by written advance directives are valuable to help loved ones, physicians, and other providers better understand and make treatment decisions consistent with patients' wishes, if the patient becomes incapacitated (5-7). Experts recommend that advance care planning discussions be held before patients face an acute health crisis and that the process continue as patients age or their clinical situation changes (3,7). However, little is known about the extent to which advance care planning is happening among very elderly community-dwellers (8-11).

We surveyed surviving community-dwelling participants of the original cohort of the Framingham Heart Study, all of whom are over age 80, about advance care planning, use of advance directives, and preferences for health care. The Framingham Heart Study is a natural place to further our understanding of very elderly adults who have been followed closely for almost 60 years, and who are now approaching the end of their natural lifespan. The original cohort participants have been followed extensively across adulthood to collect detailed medical and social information, with very little loss to follow-up. In this context, we examined the extent to which these very elderly community-dwelling adults report advance care planning; describe their preferences for care and reported use of advance directives; and examined how these factors relate to social support, illness burden, physical functioning, cognitive functioning, depressive symptoms, health perceptions, and previous health care use.

METHODS

Data Collection

The Framingham Heart Study (FHS) is a prospective observational cohort study that began in 1948 to investigate risk factors for cardiovascular disease and other health conditions. The original cohort consisted of 5,209 participants (55% women) aged 30 to 62 at entry from Framingham, Massachusetts (12). Since study inception, participants have returned every two years for a routine research examination, which include a physician-administered medical history, a medical assessment, and questionnaires administered by trained interviewers. Written informed consent was obtained from participants and the Institutional Review Board at Boston Medical Center approved the examination content.

Study Sample

Surviving original cohort participants who attended their 28th biennial examination were eligible for this study if they were cognitively intact and community-dwelling. All examinations occurred between February 4, 2004 and October 26, 2005. Overall 253 community-dwelling participants attended exam 28. however 17 participants fulfilled criteria for significant cognitive impairment or dementia after review by the FHS Dementia Study investigators (13); and therefore were ineligible to answer questions related to advance care

planning and health care preferences. Of the 236 eligible participants, 220 (93%) agreed to answer these questions. Overall, 60% of examinations for eligible participants occurred in the FHS clinic. The remainder occurred at offsite locations (e.g., private residences) often because of distance.

Advance Care Planning and Care Preferences

Twelve items were administered to assess advance care planning and preferences for care; 11 of which were administered in the SUPPORT/HELP Project and utilized in several published reports (14-16). Participants were asked whether they have talked to anyone about their wishes for medical care toward the end of life; had filled out a living will, and; completed a health care proxy. Those reporting a health care proxy were asked to identify that individual. Participants were asked whether they prefer a treatment plan focused on extending life as much as possible, even if it meant more pain and discomfort, or one focused on relieving pain and discomfort as much as possible, even if it means not living as long. Participants reported their willingness to endure certain health states including whether they would rather die. The hypothetical health states included being 1) in a great deal of pain all the time even with medications; 2) attached to a ventilator or respirator all the time; 3) fed through a tube all the time; 4) unconscious or in a coma all the time, and; 5) confused or forgetting all the time. Lastly, participants were asked where they would prefer to die.

Participants were queried about their perceptions of their own longevity and future physical functioning, factors that may influence advance care planning and care preferences. Specifically, they were asked "What do you think the chances are that you would live 12 months or more?" and "What are the chances that you will be able to take care of yourself 12 months from now?".

Due to concerns about participant burden and the potential emotional content of the questions, the interviewer assessed participants' willingness to respond to the questions and external behavior immediately following these questions. Interviewers documented whether the participant 1) stopped the interview; 2) was visibly upset or bothered by any question(s), and; 3) had difficulty understanding any question(s).

Potential Correlates

Factors previously shown to correlate with advance care planning and preferences for care in seriously ill populations are collected routinely during biennial examinations. Sociodemographic characteristics included age, sex, place of residence, and education (obtained from baseline interview). Social support was assessed by marital status, living situation, and two Berkman (17) Social Network questionnaire items, which assessed whether participants had someone to listen to them and provide emotional support.

Perceived health was assessed by *"In general, how is your health now?"*. Self-reported physical functioning was assessed by ability to perform activities of daily living (ADL) (eating, dressing, bathing, transferring, toileting, and walking about 50 yards) without human assistance (18) and ability to walk a half mile without help (19). Cognitive functioning was assessed using the Mini-mental State Exam (MMSE) [scores ≥24 indicate no cognitive impairment] (20). Presence of depressive symptoms was assessed using the Center for Epidemiologic Studies Depression Scale (CES-D) [scores ≥ 16 indicate significant depressive symptoms] (21). Use of health care services included hospitalizations, illness visit to doctor, check up by doctor, and nursing home or skilled nursing facility admission in previous 2 years. Comorbid illness was assessed by a documented history of diabetes, coronary heart disease, stroke, congestive heart failure, intermittent claudication, cancer, and hip fracture. These conditions were documented by chart review and validated by a panel of FHS investigators. Diabetes was

defined by a causal blood sugar > 200 mg/dL or self-reported use of oral hypoglycemic medications or insulin.

Statistical Analysis

All statistical analyses used SAS version 9.1 (SAS Institute, Cary, NC). We conducted bivariable analyses to describe the sample and identify differences in advance care planning and health care preferences between men and women, and to identify factors associated with self-reported use of advance directives (having a health care proxy and living will). We used multivariable logistic regression to identify factors independently associated with advanced directives after adjusting for age and sex. Because advanced directives were relatively common (>10%), we derived adjusted prevalence ratios (aPR) and 95% confidence intervals using a log binomial regression model (22).

RESULTS

Among the 220 community-dwelling respondents, 3 (1.3%) stopped before completing questions on advance care planning and preferences, and 12 (5.5%) appeared upset or bothered by one or more questions. The average age of respondents was 88 years, 67% were women, 64% were widowed, three-quarters achieved a high school education or higher, and nearly all reported a routine check-up within the past 2 years. Compared with men, women were less often married and more often living alone (Table 1). Men were more likely to have one or more comorbid conditions and to have been hospitalized than women.

We observed no difference in advance care planning and health care preferences between men and women (Table 2). Overall, 69% of respondents reported discussing their wishes for medical care at the end of life with someone, but only 17% of these respondents discussed their wishes with a physician or other health care provider. Among those who discussed their wishes for care, the vast majority talked with a family member. Almost two-thirds reported having a health care proxy, 55% a living will, 70% had either, and 41% had both. Although 80% preferred comfort over life-extending care, a greater proportion of men reported that they prefer life-extending care. Most respondents preferred to die at home.

For each health state, a greater proportion of women than men reported that they would rather die than live out their life in a particular state (Figure 1). Although 4 out of 5 respondents preferred comfort care, fewer respondents preferred death over living out their life attached to a ventilator (63%), or fed through a tube (64%) all the time. Respondents were more willing to endure symptoms of pain and confusion, with less than half indicating that they would rather die than live out their life in a great deal of pain (46%) or be forgetful or confused (45%) all of the time. In contrast, being unconscious or in a coma all the time was undesirable to most respondents, with 82% indicating that they would rather die.

Table 3 presents the proportion of respondents reporting advance directives across different factors. Having advanced directives was associated with higher education, living in a non-private residence, having someone to listen, and not having depressive symptoms. After adjusting for age and sex, only education and goals of care remained associated with having advance directives. Compared to participants with less than a high school education, those with a high school education (aPR=1.63, 95% CI [0.94, 2.81]) and more than a high school education [aPR=2.40, 95% CI (1.41, 4.07)] were more likely to have advance directives, although the former did not achieve statistical significance. Compared to participants who preferred comfort care, those preferring life-extending care were substantially less likely to have advance directives [aPR=0.32, (0.11, 0.90)], whereas participants who were unsure about their preference were indistinguishable [aPR=0.72, (0.40, 1.27)].

DISCUSSION

We studied advance care planning and health care preferences of community-dwelling participants of the Framingham Heart Study original cohort. These men and women, now aged 84 to 100 years, have been followed for over 60 years, and have made substantial contributions to our understanding of many disease processes, most notably cardiovascular disease (23). With the current study, this cohort provides important insight into our understanding of advance care planning and health care preferences of very elderly community-dwellers who are still cognitively intact and thus able to participate in discussions regarding their goals for care. The majority of respondents reported that they preferred comfort measures at the end of life and to die at home. However, in spite of the cohort's advanced age and overwhelming preference for comfort measures, when presented with specific clinical scenarios many respondents said they were willing to endure specific life-prolonging interventions (chronic ventilator and feeding tube) to avoid death. Moreover, more than half of respondents said they were willing to live out their life in a great deal of pain or being forgetful or confused rather than die.

There are several possible clinically relevant explanations for this finding. First, the framing of the questions about goals of care and hypothetical health states is very important. Patients likely do not fully understand what comfort care entails or the implications of specific life-prolonging interventions, and may not view these concepts as mutually exclusive. Moreover, patients' lack of knowledge and experience with specific medical interventions and hypothetical scenarios has been shown to cloud the relationship between goals of care and preferences for life-prolonging interventions (24). A recent study found that using video images to depict a patient living with advanced dementia had a significant impact on subjects' preferences for care by improving their understanding and ability to imagine themselves having advanced dementia (25). Prior to seeing the video, 50% of subjects preferred comfort care and 21% preferred life-prolonging care. However, after seeing the video almost 90% indicated they desired comfort care and none chose life-prolonging care. In our study, most respondents chose death over living the rest of their life in a coma or unconscious, possibly because this is a tangible state that is more readily understood. Finally, qualitative data suggest that when faced with hypothetical choices, patients place more emphasis on the outcome of an intervention (in this instance avoiding death) rather than the intervention itself (26). Nevertheless, given the inconsistency observed between reported goals of care and preferences for life-prolonging interventions and distressing health states, it is concerning that few of these very elderly respondents have discussed their wishes for end of life care with a health care provider.

More than 15 years ago, Lo and colleagues recommended that physicians talk to their elderly patients about their wishes for care (27). Most physicians believe that it is their professional responsibility to help patients with advance care planning, with over 80% of physicians reporting that they should initiate end-of-life discussions with patients (28). Therefore, it is troubling that in this very elderly cohort where nearly all respondents reported having a recent routine check-up, only 12% of all respondents reported discussing their preferences with a physician or health care provider. Yet these findings are consistent with studies suggesting that discussions with physicians about end of life care were uncommon (10,11,29). We did find it encouraging that the vast majority of respondents had someone to provide emotional support, help make difficult decisions, and listen to them, and many reported that they had discussed their wishes for care with family members. Nonetheless, our findings highlight the need for physicians to have these discussions with very elderly patients in order to help their patients make informed decisions rooted in their values.

Our findings were remarkably consistent with those of seriously ill hospitalized patients aged 80 and older enrolled in the Hospitalized Elderly Longitudinal Project (HELP) (15). Using the same set of questions asked in our study, HELP found that 73-78% of very elderly adults

preferred comfort care within 3 to 6 months prior to death. HELP also found that fewer patients would choose death rather than endure a lifetime of pain (48%) and confusion (35%) or be attached to a ventilator (70%) or feeding tube (50%) (15).

Few studies have actually examined advance care planning among community-dwelling elders (8-11). We found that two-thirds of very elderly community-dwellers reported having a health care proxy and about half had a living will. In contrast, a study of managed care patients age 80 and older found that only 27% had been asked about their end of life preferences and that only 46% had advance directives documented in their chart (10). In a multi-ethnic sample of younger community-dwelling elders, Morrison and Meier found that only 35% had a health care proxy (8).

There is some evidence that discussions about end of life care are helpful to patients. Patients who discussed end of life care with their physicians reported less fear and anxiety, a better understanding of their options for care, and a greater ability to make decisions and influence their medical care (30,31). Data suggest that completion of advance directives were associated with greater satisfaction (30,32), greater hospice use, and fewer concerns about communication (33). Despite national trends toward shared-decision making, one-third of the general adult population moderately or strongly agreed that they would prefer to leave decision-making to their physician, with older adults and those in poorer health less likely to want to participate in decision-making (34,35). Moreover, few desire tight control over medical decisions if unable to make their own decisions (36), and many community-dwelling elders trust that physicians would make the right care decisions should they become very sick (8).

Because many very elderly patients are clinically complex, advance care planning should not wait until patients face an acute medical crisis such as being hospitalized with life-threatening illness (4,7). Studies show that patients are open and willing to discuss advance care planning, but generally they would prefer that their physician raise the topic (8,37). Even though most physicians believe it is their responsibility to have end of life discussions (28), some may be reluctant since a potentially time-consuming discussion may take away from other pressing clinical matters (38). Ideally, providers should begin talking to their elderly patients about advance care planning early and have ongoing discussions, especially with changes in health status (3). This will introduce concepts of end of life care while elderly patients are able to participate in the decision-making process and allow patients to re-evaluate their values as they age or experience changes in their clinical situation.

Advance directives assume that patients can anticipate their preferences for care for hypothetical future heath states (8,39). However, there is little evidence that decisions patients make when relatively healthy can predict treatment choices when death is imminent (39). Although the vast majority of very elderly participants in our study could state their health care preferences and were not emotionally upset by the end-of-life content, some had difficulty. We found that higher education was strongly associated with advanced directives in the very elderly. Because an estimated 27% of older adults have below basic document literacy (40), health literacy may be an important and unrecognized barrier to completing advance directives. Moreover, treatment preferences appear to be only moderately stable over time, but preferences to refuse life-prolonging treatment tend to be more stable than preferences to receive life-prolonging treatment (41). Given the advanced age of our sample and that the majority preferred comfort measures — it is likely that their goals of care would be relatively stable, but it is unclear how their preferences for specific life-prolonging interventions and distressing health states would change.

Our findings should be interpreted within the context of important limitations. These findings reflect views of non-Hispanic White community-dwelling elders, most of whom still reside in

the Northeast, and may not be representative of elders from racial and ethnic minority groups or other geographic regions. In particular, studies have consistently shown that African American patients are more likely to undergo aggressive care, to prefer life-sustaining treatment, to want cardiopulmonary resuscitation, and to die in a hospital (11,32,42-48). African Americans are less likely to have discussions about life-sustaining treatments with physicians (49) and are much less accepting of advanced care planning. African Americans and Hispanic Americans have advanced directives less often than their White counterparts (44,50-53). In addition, frail older ethnic minorities are less likely to self-express their own health care choices than their White counterparts (54) Next, we were unable to look in-depth at any one topic or explore reasons why few very elderly community-dwellers discuss wishes for care with their providers. Finally, we relied on participants' self-report. It is possible that providers have addressed end of life issues with some participants, but participants did not recall or perceived the discussion differently. Although such information is subject to recall bias and participants' perceptions, perhaps it speaks to the quality of the patient-provider communication.

In summary, many very elderly adults still reside in community settings and are capable of making treatment decisions near the end of life. Most elders were able to state their preferences, and few appeared upset or bothered when asked about end of life care. Although the majority of participants prefer comfort measures at the end of life, many indicated that they would endure life-prolonging interventions and distressing health states to avoid death. The reasons for this discrepancy are likely to be multifaceted. Yet, very few respondents reported having discussed their preferences with their providers. Future research should develop and test interventions to improve rates of end of life discussions in very elderly community-dwelling populations. These efforts should involve family members whenever possible. Discussions among the physician-patient-family triad may facilitate unified decision-making that is informed and rooted in patients' values. Increased family involvement may help bridge outpatient and inpatient settings so that patients receive care that is consistent with their preferences.

Acknowledgements

The authors thank Sidney and Dorothy Stogel for providing funding for this study and acknowledge the support of Richard A. Parker, MD and Russell S. Phillips, MD. The National Heart, Lung, and Blood Institute's (NHLBI) Framingham Heart Study is supported by contract number N01-HC-25195. Dr. Kelly-Hayes was also supported by grants from the National Institute on Aging (R01 AG08122, Epidemiology of Dementia in the Framingham Study) and National Institute of Neurological Disorders and Stroke (R01 NS17950, Precursors of Stroke Incidence and Prognosis). The funding agency had no role in the design, methods, subject recruitment, data collection, analysis or preparation of the manuscript.

REFERENCES

1. Field, M.; Cassel, CK., editors. Approaching Death: Improving Care at the End of Life. National Academy Press; Washington DC: 1997.
2. AGS Ethics Committee. The care of dying patients: A position statement from the American Geriatrics Society. J Am Geriatr Soc 1995;43:577–8. [PubMed: 7537290]
3. Wenger NS, Rosenfeld K. Quality indicators for end-of-life care in vulnerable elders. Ann Intern Med 2001;135(8 Pt 2):677–85. [PubMed: 11601950]
4. Emanuel LL, von Gunten CF, Ferris FD. Advance care planning. Arch Fam Med 2000;9(10):1181–7. [PubMed: 11115227]
5. Emanuel LL, Danis M, Pearlman RA, Singer PA. Advance care planning as a process: structuring the discussions in practice. J Am Geriatr Soc 1995;43(4):440–6. [PubMed: 7706637]
6. Gillick MR. Advance care planning. N Engl J Med 2004;350(1):7–8. [PubMed: 14702421]
7. Teno JM, Lynn J. Putting advance-care planning into action. J Clin Ethics 1996;7(3):205–13. [PubMed: 8981191]

8. Morrison RS, Meier DE. High rates of advance care planning in New York City's elderly population. Arch Intern Med 2004;164(22):2421–6. [PubMed: 15596631]
9. Hopp FP. Preferences for surrogate decision makers, informal communication, and advance directives among community-dwelling elders: results from a national study. Gerontologist 2000;40(4):449–57. [PubMed: 10961034]
10. Gordon NP, Shade SB. Advance directives are more likely among seniors asked about end-of-life care preferences. Arch Intern Med 1999;159(7):701–4. [PubMed: 10218749]
11. Hanson LC, Earp JA, Garrett J, Menon M, Danis M. Community physicians who provide terminal care. Arch Intern Med 1999;159(10):1133–8. [PubMed: 10335692]
12. Dawber, T. The Framingham Study: The Epidemiology of Atherosclerotic Disease. Harvard University Press; Cambridge, MA: 1980.
13. Seshadri S, Wolf PA, Beiser A, et al. Lifetime risk of dementia and Alzheimer's disease. The impact of mortality on risk estimates in the Framingham Study. Neurology 1997;49(6):1498–504. [PubMed: 9409336]
14. McCarthy EP, Phillips RS, Zhong Z, Drews RE, Lynn J. Dying with cancer: patients' function, symptoms, and care preferences as death approaches. J Am Geriatr Soc 2000;48(5 Suppl):S110–21. [PubMed: 10809464]
15. Somogyi-Zalud E, Zhong Z, Lynn J, Hamel MB. Elderly persons' last six months of life: findings from the Hospitalized Elderly Longitudinal Project. J Am Geriatr Soc 2000;48(5 Suppl):S131–9. [PubMed: 10809466]
16. Weeks JC, Cook EF, O'Day SJ, et al. Relationship between cancer patients' predictions of prognosis and their treatment preferences. JAMA 1998;279(21):1709–14. [PubMed: 9624023]
17. Berkman LF, Syme SL. Social networks, host resistance, and mortality: a nine-year follow-up study of Alameda County residents. Am J Epidemiol 1979;109(2):186–204. [PubMed: 425958]
18. Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of Illness in the Aged. the Index of Adl: a Standardized Measure of Biological and Psychosocial Function. JAMA 1963;185:914–9. [PubMed: 14044222]
19. Rosow I, Breslau N. A Guttman health scale for the aged. J Gerontol 1966;21(4):556–9. [PubMed: 5918309]
20. 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(3):189–98. [PubMed: 1202204]
21. Radloff LS. The CES-D scale: a self-report depression scale for research in the general population. Appl Psychol Meas 1977;1:385–401.
22. Spiegelman D, Hertzmark E. Easy SAS calculations for risk or prevalence ratios and differences. Am J Epidemiol 2005;162(3):199–200. [PubMed: 15987728]
23. Framingham Heart Study. [Accessed January 8, 2008]. http://www.framinghamheartstudy.org/biblio/index.html
24. Brett AS. Limitations of listing specific medical interventions in advance directives. JAMA 1991;266 (6):825–8. [PubMed: 1865521]
25. Volandes AE, Lehmann LS, Cook EF, Shaykevich S, Abbo ED, Gillick MR. Using video images of dementia in advance care planning. Arch Intern Med 2007;167(8):828–33. [PubMed: 17452547]
26. Rosenfeld KE, Wenger NS, Kagawa-Singer M. End-of-life decision making: a qualitative study of elderly individuals. J Gen Intern Med 2000;15(9):620–5. [PubMed: 11029675]
27. Lo B, Rouse F, Dornbrand L. Family decision making on trial. Who decides for incompetent patients? N Engl J Med 1990;322(17):1228–32. [PubMed: 2109196]
28. Markson L, Clark J, Glantz L, Lamberton V, Kern D, Stollerman G. The doctor's role in discussing advance preferences for end-of-life care: perceptions of physicians practicing in the VA. J Am Geriatr Soc 1997;45(4):399–406. [PubMed: 9100706]
29. McDonald DD, Deloge JA, Joslin N, et al. Communicating end-of-life preferences. West J Nurs Res 2003;25(6):652–66. [PubMed: 14528616]discussion 667-75
30. Smucker WD, Ditto PH, Moore KA, Druley JA, Danks JH, Townsend A. Elderly outpatients respond favorably to a physician-initiated advance directive discussion. J Am Board Fam Pract 1993;6(5): 473–82. [PubMed: 8213238]

31. Ditto PH, Danks JH, Smucker WD, et al. Advance directives as acts of communication: a randomized controlled trial. Arch Intern Med 2001;161(3):421–30. [PubMed: 11176768]
32. Tierney WM, Dexter PR, Gramelspacher GP, Perkins AJ, Zhou XH, Wolinsky FD. The effect of discussions about advance directives on patients' satisfaction with primary care. J Gen Intern Med 2001;16(1):32–40. [PubMed: 11251748]
33. Teno JM, Gruneir A, Schwartz Z, Nanda A, Wetle T. Association between advance directives and quality of end-of-life care: a national study. J Am Geriatr Soc 2007;55(2):189–94. [PubMed: 17302654]
34. Levinson W, Kao A, Kuby A, Thisted RA. Not all patients want to participate in decision making. A national study of public preferences. J Gen Intern Med 2005;20(6):531–5. [PubMed: 15987329]
35. Demoratz MJ. Advance directives: getting patients to complete them before they need them. Case Manager 2005;16(1):61–3. [PubMed: 15678027]
36. Hawkins NA, Ditto PH, Danks JH, Smucker WD. Micromanaging death: process preferences, values, and goals in end-of-life medical decision making. Gerontologist 2005;45(1):107–17. [PubMed: 15695421]
37. Emanuel LL, Barry MJ, Stoeckle JD, Ettelson LM, Emanuel EJ. Advance directives for medical care--a case for greater use. N Engl J Med 1991;324(13):889–95. [PubMed: 2000111]
38. Ramsaroop SD, Reid MC, Adelman RD. Completing an advance directive in the primary care setting: what do we need for success? J Am Geriatr Soc 2007;55(2):277–83. [PubMed: 17302667]
39. Meier DE, Morrison RS. Autonomy reconsidered. N Engl J Med 2002;346(14):1087–9. [PubMed: 11932479]
40. Marcus EN. The silent epidemic--the health effects of illiteracy. N Engl J Med 2006;355(4):339–41. [PubMed: 16870912]
41. Ditto PH, Smucker WD, Danks JH, et al. Stability of older adults' preferences for life-sustaining medical treatment. Health Psychol 2003;22(6):605–15. [PubMed: 14640858]
42. Shepardson LB, Gordon HS, Ibrahim SA, Harper DL, Rosenthal GE. Racial variation in the use of do-not-resuscitate orders. J Gen Intern Med 1999;14(1):15–20. [PubMed: 9893086]
43. McKinley ED, Garrett JM, Evans AT, Danis M. Differences in end-of-life decision making among black and white ambulatory cancer patients. J Gen Intern Med 1996;11(11):651–6. [PubMed: 9120650]
44. Blackhall LJ, Frank G, Murphy ST, Michel V, Palmer JM, Azen SP. Ethnicity and attitudes towards life sustaining technology. Soc Sci Med 1999;48(12):1779–89. [PubMed: 10405016]
45. Phillips RS, Wenger NS, Teno J, et al. Choices of seriously ill patients about cardiopulmonary resuscitation: correlates and outcomes. SUPPORT Investigators. Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatments. Am J Med 1996;100(2):128–37. [PubMed: 8629646]
46. Rosenfeld KE, Wenger NS, Phillips RS, et al. Factors associated with change in resuscitation preference of seriously ill patients. The SUPPORT Investigators. Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatments. Arch Intern Med 1996;156(14):1558–64. [PubMed: 8687264]
47. Goodlin SJ, Zhong Z, Lynn J, et al. Factors associated with use of cardiopulmonary resuscitation in seriously ill hospitalized adults. JAMA 1999;282(24):2333–9. [PubMed: 10612321]
48. Eleazer GP, Hornung CA, Egbert CB, et al. The relationship between ethnicity and advance directives in a frail older population. J Am Geriatr Soc 1996;44(8):938–43. [PubMed: 8708304]
49. Haas JS, Weissman JS, Cleary PD, et al. Discussion of preferences for life-sustaining care by persons with AIDS. Predictors of failure in patient-physician communication. Arch Intern Med 1993;153 (10):1241–8. [PubMed: 8494476]
50. Morrison RS, Zayas LH, Mulvihill M, Baskin SA, Meier DE. Barriers to completion of health care proxies: an examination of ethnic differences. Arch Intern Med 1998;158(22):2493–7. [PubMed: 9855388]
51. Degenholtz HB, Arnold RA, Meisel A, Lave JR. Persistence of racial disparities in advance care plan documents among nursing home residents. J Am Geriatr Soc 2002;50(2):378–81. [PubMed: 12028224]

52. Degenholtz HB, Thomas SB, Miller MJ. Race and the intensive care unit: disparities and preferences for end-of-life care. Crit Care Med 2003;31(5 Suppl):S373–8. [PubMed: 12771586]

53. Kiely DK, Mitchell SL, Marlow A, Murphy KM, Morris JN. Racial and state differences in the designation of advance directives in nursing home residents. J Am Geriatr Soc 2001;49(10):1346–52. [PubMed: 11890494]

54. Hornung CA, Eleazer GP, Strothers HS 3rd, et al. Ethnicity and decision-makers in a group of frail older people. J Am Geriatr Soc 1998;46(3):280–6. [PubMed: 9514372]

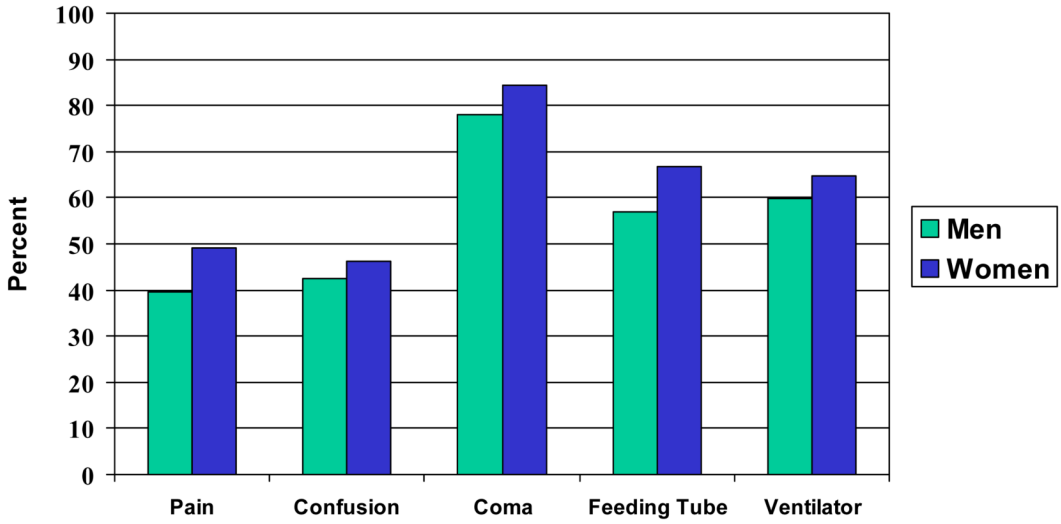


Figure 1.
Preferences for Future Health States among Community-Dwelling Elders (n = 220)
Respondent would rather die than spend all of the time in:

Table 1

Characteristics of Community-Dwelling Participants of the Framingham Heart Study by Sex (n=220)

	Women (n=147) n (%)	Men (n=73) n (%)	p-value
Demographic Characteristics			
Age at interview, mean ± SD	88.3 ± 3.4	87.8 ± 3.1	.277
Education			
Less than high school graduate	32 (22)	19 (26)	
High school graduate	64 (44)	29 (40)	
More than high school	48 (33)	26 (34)	.789
Place of residence			
Private residence	111 (76)	58 (79)	
Other (e.g. assisted living)	36 (24)	15 (21)	.514
Social Support			
Marital status			
Married	21 (14)	41 (58)	
Widow	111 (76)	29 (41)	
Other	15 (10)	1 (1)	<.001
Living situation			
Lives alone	87 (60)	22 (31)	
Lives with someone	59 (40)	50 (69)	<.001
Has someone to listen			
None of the time	7 (5)	1 (2)	
Some of the time	43 (30)	22 (32)	
All of the time	91 (65)	45 (66)	.536
Has someone to provide emotional support and help make difficult decisions			
None of the time	6 (4)	3 (4)	
Some of the time	20 (14)	15 (22)	
All of the time	114 (81)	52 (74)	.353
Clinical Characteristics			
History of			
Cancer	39 (27)	26 (36)	.164
Coronary heart disease	38 (26)	29 (40)	.035
Stroke	10 (7)	17 (23)	<.001
Intermittent claudication	12 (9)	12 (16)	.095
Congestive heart failure	14 (10)	9 (12)	.522
Diabetes	18 (12)	11 (15)	.560
Hip fracture	17 (12)	2 (3)	.039
Number of comorbid illnesses			
None	56 (38)	13 (18)	
1	48 (33)	26 (36)	
2	29 (20)	23 (31)	
3 or more	14 (9)	11 (15)	.002

	Women (n=147) n (%)	Men (n=73) n (%)	p-value
Perceived health			
Excellent	25 (17)	12 (17)	
Good	95 (66)	45 (62)	
Fair/poor	24 (17)	15 (21)	.571
Perceived chance of taking care of self in 12 months			
90% or better	81 (55)	39 (54)	
about 75%	19 (13)	19 (26)	
50% or worse	37 (26)	12 (17)	
Unsure	9 (6)	2 (3)	.053
Perceived chance of living 12 months or longer			
90% or better	85 (55)	43 (60)	
about 75%	11 (7)	11 (15)	
50% or worse	39(27)	13 (18)	
Unsure	12 (8)	5 (7)	.211
CES-D* Score ≥ 16, indicating significant depressive symptoms			
No	119 (82)	64 (89)	
Yes	27 (18)	8 (11)	.176
Cognitive status, MMSE Score††			
< 24	132 (90)	66 (90)	
≥ 24	15 (10)	7 (10)	.886
ADL† dependencies			
None	128 (87)	65 (89)	
1 or more	19 (13)	8 (11)	..676
Uses assistive device to perform ADL†			
No	49 (33)	32 (44)	
Yes	98 (67)	41 (56)	.120
Ability to walk a half mile without help	57 (39)	20 (28)	
Able	63 (44)	43 (60)	
Unable	25 (17)	9 (12)	.078
Doesn't do			
Health care use since last exam			
Had check up by doctor	141 (97)	70 (96)	1.00
Illness visit to doctor			
No visit	96 (65)	39 (53)	
1	29 (20)	24 (33)	
2 or more	22 (15)	10 (14)	.314
Hospitalization			
None	97 (66)	35 (48)	

	Women (n=147) n (%)	Men (n=73) n (%)	p-value
1	28 (19)	24 (33)	
2 or more	22 (15)	14 (19)	.041
Had nursing home stay	18 (12)	9 (12)	.986

Number of observations with missing data — education (n=3); marital status (n=2); living alone (n=2); having someone to listen to you (n=11); having someone to provide emotional support and help make difficult decisions (n=10); perceived ability to take care of self (n=2); perception of living 12 months or longer (n=1); CES-D (n=2); ADLs (n=1); ability to walk a half mile without help (n=3); Interim check up by doctor (n=1).

* CES-D = Center for Epidemiologic Studies Depression Scale.

†† MMSE = Mini-Mental Status Exam

† ADL = Activities of daily living.

Table 2

Advance Care Planning and Health Care Preferences among Community-Dwelling Elders (n=220)

	All Participants[†] n (%)	Women (n=147) n (%)	Men (n=73) n (%)
Discussed wishes for end of life care			
No	69 (31)	42 (29)	27 (37)
Yes	151 (69)	105 (71)	46 (63)
Among those who have discussed wishes, percent who discussed with	[n=151]	[n=105]	[n=46]
Family member	138 (91)	96 (91)	42 (91)
Physician/health care provider	25 (17)	17 (16)	8 (17)
Attorney	25 (17)	15 (14)	10 (21)
Friend	17 (11)	15 (14)	2 (4)
Clergy	5 (3)	3 (3)	2 (4)
Other	5 (3)	5 (5)	0 (0)
Completed advance directives (has health care proxy and living will)	91 (41)	60 (41)	31 (43)
Has heath care proxy	136 (66)	92 (67)	44 (62)
Has Living Will	116 (55)	75 (54)	41 (57)
Preferences for Goals of Care			
Extend life as much as possible	18 (8)	8 (5)	10 (14)
Focused on comfort/pain relief	174 (80)	122 (84)	52 (72)
Unsure	26 (12)	16 (11)	10 (14)
Preference for Place of Death			
Home	139 (71)	93 (69)	46 (75)
Hospital	23 (12)	16 (12)	7 (12)
Hospice	14 (7)	12 (9)	2 (3)
Nursing home	2 (1)	2 (2)	0 (0)
Other	17 (9)	11 (8)	6 (10)

[†]There were no statistically significant differences by sex. Number of observations with missing data — health care proxy (n=12); living will (n=8); preference for goals of care (n=2); preference for place of death (n=25).

Table 3

Proportion of Community-Dwelling Elders with Advance Directives Across Selected Characteristics (n=220)

	Number with Characteristic (n)	% with Advance Directives*
Demographic Characteristics		
Sex		
Female	147	40.8
Male	73	42.5
Education†		
Less than high school graduate	51	23.5
High school graduate	93	40.9
More than high school	73	54.8
Place of residence†		
Private residence	169	37.9
Other (e.g. assisted living)	51	52.9
Social Support		
Marital status		
Married	62	46.8
Widow	140	41.4
Other	16	25.0
Lives situation		
Lives alone	109	45.0
Lives with someone	109	36.7
Has someone to listen†		
None of the time	8	12.5
Some of the time	65	32.3
All of the time	136	49.3
Has someone to provide emotional support and help make difficult decisions		
None of the time	9	22.2
Some of the time	35	34.3
All of the time	166	45.2
Clinical Characteristics		
Number of comorbid illnesses		
None	69	44.9
1	74	39.2
2	52	42.3
3 or more	25	36.0
Perceived health status		
Excellent	37	43.2
Good	140	41.4
Fair/poor	39	41.0
Perceived chance of living 12 months or longer		
90% or better	128	45.3

	Number with Characteristic (n)	% with Advance Directives*
about 75%	22	31.8
50% or worse	52	36.5
Unsure	17	41.2
Perceived chance of taking care of self in 12 months		
90% or better	120	48.3
about 75%	38	36.8
50% or worse	49	28.6
Unsure	11	36.4
CES-D+ Score ≥ 16, indicating significant depressive symptoms†		
No	183	44.8
Yes	35	25.7
Cognitive status, MMSE Score††		
< 24	198	42.4
≥ 24	22	31.8
ADL§ dependencies		
None	193	43.0
1 or more	27	29.6
Uses assistive device to perform ADL§		
No	81	45.7
Yes	139	38.9
Ability to walk a half mile without help		
Able	106	49.0
Unable	77	33.8
Doesn't do	34	32.3
Health care use since last exam		
Check up by doctor		
No	8	37.5
Yes	211	41.7
Illness visit to doctor		
No visit	135	37.8
1	53	52.8
2 or more	12	37.5
Hospitalization		
None	132	42.4
1	52	36.5
2 or more	36	44.4
Nursing home stay		
No	193	43.5
Yes	27	25.9

Number of observations with missing data — education (n=3); marital status (n=2); perceived health (n=4); living alone (n=2); having someone to listen to you (n=11); having someone to provide emotional support and help make difficult decisions (n=10); perceived health status (n=4); perceived ability to take care of self (n=13); perception of living 12 months or longer (n=18); CES-D (n=2); ADLs (n=1); Interim check up by doctor (n=1).

[*] Percentages refer to proportion of participants with a particular characteristic who had an advance directive.

[†] p-value <0.05.

[+] CES-D = Center for Epidemiologic Studies Depression Scale.

[††] MMSE = Mini-Mental Status Exam

[§] ADL = Activities of daily living.

Patient Risks, Outcomes, and Costs of Voluntary HIV Testing at Five Testing Sites Within a Medical Center

Supriya D. Mehta, PhD, MHS[a]
Jonathan Hall, MM, LADC-1[b]
Jeffrey L. Greenwald, MD[c]
Kevin Cranston, MD[d]
Paul R. Skolnik, MD[b,e]

SYNOPSIS

Objectives. The Centers for Disease Control and Prevention (CDC) recommends offering human immunodeficiency virus (HIV) testing to all patients in all high HIV-prevalence clinical settings. We evaluated programmatic aspects of HIV testing across multiple clinical settings within a single medical center.

Methods. We analyzed programmatic data of HIV testing in the Urgent Care Center (UCC), inpatient floors, outpatient primary care, a non-clinical Drop-In Center, and Emergency Department (ED). HIV testing was by oral mucosal transudate, venous blood samples, or rapid testing fingersticks, with Western blot confirmation. We compared the sociodemographics and behavioral risks of individuals undergoing HIV testing across the five sites and estimated costs per person tested and per HIV-positive test result.

Results. From 2002 to 2004, 16,750 HIV tests were conducted, with 229 (1.4%) previously unreported HIV infections diagnosed among 16,696 valid test results. HIV-positive prevalence was 1.5% for the UCC, 1.5% at the Drop-In Center, 1.4% for primary care, 1.2% for inpatient, and 0.6% in the ED. Behavioral risks were most prevalent in the UCC and the Drop-In Center. The cost per test was lowest in the UCC and highest in the Drop-In Center. The cost per previously unreported HIV infection was lowest in the UCC ($1,980) and highest in the ED ($9,724).

Conclusions. Although a significant number of HIV infections were identified, the number of tests performed represents <10% of all clinical visits. Due to personnel and time constraints, offering HIV testing to patients hierarchically in some settings of a high-volume medical center merits evaluation.

[a]Department of Epidemiology and Biostatistics, University of Illinois Chicago, Chicago, IL

[b]Department of Medicine, Center for HIV/AIDS Care and Research, Boston Medical Center, Boston, MA

[c]Hospital Medicine Unit, Department of Medicine, Boston Medical Center and Boston University School of Medicine, Boston, MA

[d]HIV/AIDS Bureau, Massachusetts Department of Public Health, Boston, MA

[e]Division of Infectious Diseases, Department of Medicine, Boston University School of Medicine, Boston, MA

Address correspondence to: Supriya D. Mehta, PhD, MHS, Department of Epidemiology and Biostatistics, University of Illinois, 958 SPHPI, M/C 923, 1603 W. Taylor St., Chicago, IL 60612; tel. 312-413-7485; fax 312-996-0064; e-mail <supriyad@uic.edu>.

Of the estimated 1 million people living with human immunodeficiency virus (HIV)/acquired immunodeficiency syndrome (AIDS) in the U.S.,[1] approximately 25% are unaware of their status.[2] In 2003, the Centers for Disease Control and Prevention (CDC) published *Advancing HIV Prevention: New Strategies for a Changing Epidemic*, augmenting previous recommendations "to include offering HIV testing to all patients in all high HIV-prevalence clinical settings and to those with risks for HIV in low HIV-prevalence clinical settings."[3] This strategy also supports simplifying HIV pretest prevention counseling and recommends opt-out HIV screening for all patients in all health-care settings.[4]

Routine screening can be advantageous compared with targeted- or self-referred screening by increasing the number of cases detected.[3,5–8] With the updated 2006 CDC recommendations advocating the elimination of specific consent for HIV tests, utilizing an opt-out protocol at clinical sites, and routinely testing for HIV for all individuals without a known HIV serostatus,[4] questions remain about the feasibility of such broad-based recommendations. Implementing such a plan will require a high degree of health-care provider acceptance, improved reimbursement, and recognition of complicated features in the process of HIV testing, most notably counseling for HIV-positive patients and linkage to care.[9]

Studies of HIV testing among Emergency Department (ED), inpatient, and urgent care patients shows the prevalence of previously undiagnosed HIV through routinely offered and targeted testing ranges from 2.0% to 4.0%.[5,10–15] None of these studies directly compared characteristics of people tested and outcomes of testing across different clinical sites. This information may help develop and refine HIV testing strategies within a hospital system, with the eventual goal of maximizing resource allocation in terms of numbers of HIV cases detected in relation to the availability of personnel time and cost.

The goals of this analysis were to compare across multiple sites within a single hospital system: (*1*) HIV test results, and patient demographics and behavioral risks and (*2*) the cost per HIV test conducted and HIV infection diagnosed.

METHODS

Study setting

Boston Medical Center (BMC) is a 547-bed academic affiliate of Boston University (BU). Approximately 50% of BMC patients are uninsured or have Medicaid. In 2004, there were approximately 28,000 hospital admissions, 55,000 outpatient primary care visits, 29,000 urgent care visits, 97,000 adult ED visits, and 28,000 pediatric ED visits.

This analysis examines counselor-based, voluntary opt-in HIV testing from January 2002 through December 2004 in five sites: the Urgent Care Center (UCC), the Drop-In Center, inpatient floors, the primary care clinic (PCC), and the ED. Patients in the UCC, inpatient, and PCC were offered confidential testing; patients in the Drop-In Center were offered anonymous or confidential testing; and patients in the ED were offered only anonymous testing. HIV counseling and testing in all sites but the ED was funded under a contract with the Massachusetts Department of Public Health (MDPH). Excluding the ED, testing was conducted Monday through Friday from 8:30 a.m. to 5 p.m. by six full-time counselors. HIV testing was offered in the ED from November 2003 to May 2004, as part of a research study approved by the BU Medical Center Institutional Review Board, evaluating routinely recommended HIV and sexually transmitted disease (STD) testing.[16] Two full-time counselors provided testing Monday through Friday from 10 a.m. to 12 a.m. in the adult ED, and one full-time counselor provided testing from 10 a.m. to 10 p.m. in the pediatric ED. Counselors approached and offered HIV testing to English- and Spanish-speaking ED patients aged 15 to 54 systematically, by going from bed to bed in a consecutive manner. Testing was offered routinely based on age, language, and order of approach; risk was not assessed prior to offering testing. ED patients who reported their HIV status to be positive were not tested.

The Drop-In Center is a non-clinical site where clients are self-referred for HIV testing. There is no pre-screening of risk; all clients are offered testing. Intake ends one hour prior to the center's closing time. Individuals who present after intake hours end are asked to return at another time. In the UCC, a hierarchical screening method was employed due to the high volume of patients. Patients were approached in examination rooms after triage and the counselors monitored the examination room board to first identify patients who presented with (in descending order of priority): infectious diseases, fever, drug or alcohol use, incarceration, STD-related complaint, or homelessness. On the inpatient floors, patients were approached in their rooms and there were also physician referrals. A hierarchical approach was not applied on inpatient floors and testing was offered prior to risk assessment. In the outpatient PCC, patients were tested mostly by physician referral. HIV testing in the UCC, Drop-In Center, inpatient floors, and PCC occurred among patients aged 15 years and older with no upper age limit.

HIV testing

From 2002 to 2003, patients gave a venous blood sample or an oral mucosal transudate (OMT) sample to be tested for HIV by enzyme-linked immunosorbent assay (ELISA), with Western blot confirmation through the MDPH State Laboratory Institute, according to manufacturer's directions (OraSure®, Orasure Technologies Inc., Bethlehem, Pennsylvania). All ED patients were tested for HIV via OMT sample. Beginning in 2004, patients were offered rapid HIV testing in the UCC, Drop-In Center, PCC, and inpatient floors. Blood samples obtained by fingerstick were tested for HIV by OraQuick® Rapid HIV-1 Antibody Test (Orasure Technologies, Inc.). Reactive rapid HIV results were retested for HIV by both ELISA and Western blot confirmation (the latter done irrespective of ELISA results) through the MDPH State Laboratory Institute.

Programmatic costs

The goal of the programmatic cost comparison across sites was to determine the relative rather than the actual costs. To determine the cost of the programs relative to each other, rather than the actual costs, we estimated the personnel costs per specimen tested and per previously unreported HIV infected person detected. This cost comparison provides an understanding of the costs associated with personnel effort and the prevalence of HIV-positive tests. Throughout the medical center, including the ED, HIV counseling and testing supplies and specimen processing were provided by the MDPH at no charge. Therefore, material costs are not included in the programmatic costs estimation.

From 2002 through 2004, there were five full-time equivalent employees (FTEEs) conducting HIV counseling and testing (excluding the ED): 1.5 FTEEs in the UCC, 1.5 FTEEs in the Drop-In Center, 1.5 FTEEs in inpatient, and 0.5 FTEEs in the PCC. From November 2003 through May 2004, there were two FTEEs in the adult ED and one FTEE in the pediatric ED. For personnel costs, each FTEE was assumed: (*1*) to have an annual salary of $35,000 (current starting salary for HIV counselors) multiplied by an institutional fringe benefit rate of 27%, and (*2*) to work five days a week for 47 weeks (holidays plus sick days = five weeks).

Data collection and analysis

Demographic and behavioral data were collected using the MDPH HIV Counseling and Testing Data Collection form at the time of HIV testing. Data were recorded through personal interviews. Demographic and behavioral data were linked to laboratory results by unique identifiers. HIV test results were dichotomized as HIV-negative (non-reactive ELISA) and HIV-positive (reactive ELISA and Western blot); we used the ELISA and Western blot confirmed results of rapid test results for HIV-negative (non-reactive ELISA, non-reactive Western blot) and HIV-positive (reactive ELISA and reactive Western blot). Data represent the number of tests conducted rather than the number of individuals, and therefore are likely to include some patients who were tested more than once. Identifying information is not recorded in the database, thus it is not possible to identify repeat visits.

Descriptive analyses include frequencies of patient demographics and behavioral risks overall and by site. Pearson's Chi-square test assessed statistical significance of differences in patient characteristics by testing site and in HIV prevalence by patient characteristics. Point prevalence and binomial exact 95% confidence intervals were calculated for HIV-positive test results by site and patient characteristics, to allow observation of within-category differences. Logistic regression was used to calculate univariable odds ratios. Data were analyzed using Stata® software.[17]

RESULTS

From January 2002 to December 2004, there were 17,594 pretest counseling sessions and 16,750 HIV tests conducted. Almost 40% of tests took place in the UCC (n=6,632), 25% from inpatient floors (n=4,253), 17% at the Drop-In Center (n=2,920), 9% from the PCC (n=1,516), and 9% in the ED (n=1,427). Among 16,750 HIV tests, 258 (1.5%) were positive. Of the 258 patients who tested HIV positive, 29 reported a previous HIV-positive test result in their pretest counseling assessment and were excluded from analysis. (Patients who report their HIV status as positive in the UCC, PCC, inpatient floors, or Drop-In Center, but do not have any existing documentation, are tested to provide documentation and facilitate entry into care.) There were 15 indeterminate results and 10 unusable specimens (nine were quantity not sufficient and one was hemolyzed) that were excluded from analysis. Thus, 229 (1.37%; 95% confidence interval [CI] 1.20, 1.56) of 16,696 patients were newly diagnosed with HIV. HIV test results were from 9,404 (56.3%) OMT samples, 1,871 (11.2%) venous blood samples, and 5,421 (32.5%) rapid testing fingersticks (all conducted in 2004).

Among all patients tested, 59% of patients were male, the median age was 36 years, 39% were black, 23% were white, 19% were Hispanic, and 8% were Haitian (Table 1). HIV prevalence was similar in the UCC (1.5%), Drop-In Center (1.5%), inpatient floors (1.4%), and PCC (1.2%), and lower in the ED (0.6%).

Table 1. Overall patient demographics and HIV results by site of HIV test, 2002–2004

	Total	*UCC*	*Drop-In Center*	*Inpatient*	*PCC*	*ED*	
	N *(percent)*	N *(percent)*	N *(percent)*	N *(percent)*	N *(percent)*	N *(percent)*	P-*value*
Age (in years)[a]							
15–24	2,930 (18.0)	1,085 (16.6)	512 (19.0)	495 (11.8)	248 (16.7)	590 (42.4)	
25–39	6,953 (42.6)	3,316 (50.6)	1,333 (49.5)	1,227 (29.3)	635 (42.3)	442 (31.8)	
40–54	5,202 (30.8)	1,775 (27.1)	740 (27.5)	1,688 (40.3)	457 (30.8)	359 (25.8)	
≥55	1,403 (8.6)	375 (5.7)	107 (4.0)	775 (18.5)	146 (9.8)	0 (0.0)	<0.001
Gender							
Female	6,776 (41.0)	2,371 (36.1)	1,092 (38.4)	1,834 (43.6)	711 (47.5)	768 (55.0)	
Male	9,734 (59.0)	4,191 (63.9)	1,753 (61.6)	2,374 (56.4)	787 (52.5)	629 (45.0)	<0.001
Race							
White	3,838 (23.0)	1,024 (15.5)	1,141 (39.2)	1,228 (29.0)	197 (13.0)	248 (17.4)	
Black	6,420 (38.5)	2,961 (44.8)	788 (27.1)	1,599 (37.7)	419 (27.6)	653 (45.8)	
Hispanic	3,184 (19.1)	1,281 (19.5)	507 (17.4)	809 (19.1)	298 (19.7)	283 (19.8)	
Haitian	1,271 (7.6)	640 (9.7)	69 (2.4)	209 (4.9)	279 (18.4)	74 (5.2)	
Cape Verdean	478 (2.9)	238 (3.6)	41 (1.4)	97 (2.3)	54 (3.6)	48 (3.4)	
Asian	321 (1.9)	112 (1.7)	87 (3.0)	51 (1.2)	62 (4.1)	9 (0.6)	
Other, unknown	1,184 (7.1)	341 (5.2)	277 (9.5)	247 (5.8)	207 (13.7)	112 (7.9)	<0.001
Education							
<High school	3,442 (20.6)	1,536 (23.3)	339 (11.7)	929 (21.9)	281 (18.5)	357 (25.0)	
≥High school/GED	11,116 (66.6)	4,776 (72.3)	2,139 (73.5)	2,170 (51.2)	1,024 (67.6)	1,007 (70.6)	
Declined, missing	2,138 (12.8)	291 (4.4)	432 (14.9)	1,141 (26.9)	211 (13.9)	63 (4.4)	<0.001
Homeless							
No	14,826 (94.6)	6,010 (93.2)	2,268 (94.8)	3,852 (94.8)	1,468 (98.5)	1,228 (96.3)	
Yes	842 (5.4)	436 (6.8)	124 (5.2)	212 (5.2)	23 (1.5)	47 (3.7)	<0.001
Referral source							
Self	11,061 (66.3)	6,315 (95.6)	1,774 (61.0)	1,815 (42.8)	213 (14.1)	944 (66.2)	
Physician	3,924 (23.5)	215 (3.3)	192 (6.6)	2,266 (53.4)	1,251 (82.5)	0 (0.0)	
Other	1,711 (10.3)	73 (1.1)	994 (32.4)	159 (3.8)	52 (3.4)	483 (33.9)	<0.001
HIV specimen source							
OMT	9,404 (56.3)	4,203 (63.4)	1,174 (40.2)	2,512 (59.1)	110 (7.3)	1,427 (100)	
Venous blood sample	1,871 (11.2)	61 (0.9)	1,183 (40.5)	32 (0.8)	625 (41.2)	0 (0.0)	
Fingerstick for rapid test	5,421 (32.5)	2,368 (35.7)	563 (19.3)	1,709 (40.2)	781 (51.5)	0 (0.0)	<0.001
HIV results							
Negative	16,467 (98.6)	6,502 (98.5)	2,867 (98.5)	4,181 (98.6)	1,498 (98.8)	1,419 (99.4)	
Positive	229 (1.4)	101 (1.5)	43 (1.5)	59 (1.4)	18 (1.2)	8 (0.6)	0.067

[a]Testing was not conducted among patients <16 years of age in the UCC, Drop-In Center, or outpatient PCC, or among patients <15 years of age in inpatient floors and the ED. Programmatically, HIV testing was not offered to ED patients >54 years of age. Statistical comparison excludes the ED.

HIV = human immunodeficiency virus

UCC = Urgent Care Center

PCC = primary care clinic

ED = Emergency Department

GED = graduated equivalent degree

OMT = oral mucosal transudate

The proportion of male clients HIV tested was lowest in the ED (45%), while men comprised more than 60% of clients tested in the UCC and Drop-In Center. A greater proportion of ED and UCC patients were black, and a greater proportion of Drop-In Center and inpatient floor clients were white. Almost 60% of clients undergoing HIV testing on the inpatient floors were aged ≥40 years, compared with 32.8% in the UCC, 31.5% in the Drop-In Center, and 40.6% in the PCC. Behavioral risks were more prevalent among UCC and Drop-In Center patients (Table 2). Patients tested at the Drop-In Center were more likely to report same-sex intercourse, injection drug use, or sex with an HIV-positive person than patients at any other testing venue. Patients tested in the UCC were more likely to report sex under the influence of drugs or alcohol, or a history of STD in the past three years.

HIV-positive test results by patient characteristics

Among demographic factors, patients who were male, aged 25 to 54, of non-white race other than Asian, uninsured, with less than a high school education, or referred for HIV testing by a physician were more likely to test HIV positive (Table 3). The prevalence of infection was greater than 2% among patients who were Haitian, Cape Verdean, uninsured, or had less than a high school education. HIV-positive results were more prevalent among patients who reported same-sex intercourse, history of STD, sex with an HIV-positive person, inconsistent condom use, or anal-receptive and anal-insertive intercourse. Other factors by which HIV test results did not differ (results not shown) were: sex under the influence of drugs or alcohol, sex with a commercial sex partner, sex in exchange for drugs or money, homelessness, and having a casual sex partner.

Table 2. Patient risks by site of HIV test, 2002–2004[a]

	Total	*UCC*	*Drop-In Center*	*Inpatient*	*PCC*	*ED*	
	N *(percent)*	N *(percent)*	N *(percent)*	N *(percent)*	N *(percent)*	N *(percent)*	P-*value*
Sex partner gender							
Opposite gender only	14,308 (91.8)	5,979 (94.1)	2,168 (79.3)	3,628 (95.8)	1,320 (94.0)	1,213 (93.0)	
Same or both genders	1,278 (8.2)	377 (5.9)	567 (20.7)	159 (4.2)	85 (6.0)	91 (7.0)	<0.001
IDU							
No	15,372 (92.1)	6,044 (91.5)	2,474 (85.0)	4,007 (94.5)	1,481 (97.7)	1,366 (95.7)	
Yes	1,324 (7.9)	559 (8.5)	436 (15.0)	233 (5.5)	35 (2.3)	61 (4.3)	<0.001
Use of needle exchange program among IDUs							
No	795 (63.2)	359 (65.5)	262 (61.9)	116 (58.9)	25 (80.6)	33 (55.9)	
Yes	463 (36.8)	189 (35.5)	161 (38.1)	81 (41.1)	6 (19.4)	26 (44.1)	0.074
Sex under the influence of drugs or alcohol							
No	12,061 (72.2)	3,920 (59.4)	2,328 (80.0)	3,577 (84.4)	1,233 (81.3)	1,003 (70.3)	
Yes	4,635 (27.8)	2,683 (40.6)	582 (20.0)	663 (15.6)	283 (18.7)	424 (29.7)	<0.001
STD diagnosis							
No	13,696 (82.0)	4,322 (65.5)	2,592 (89.1)	4,064 (95.9)	1,436 (94.7)	1,282 (89.8)	
Yes	3,000 (18.0)	2,281 (34.5)	318 (10.9)	176 (4.1)	80 (5.3)	145 (10.2)	<0.001
Sex with an HIV-positive person	16,100 (96.4)	6,327 (95.8)	2,677 (92.0)	4,205 (99.2)	1,492 (98.4)	1,399 (98.0)	
No	596 (3.6)	276 (4.2)	233 (8.0)	35 (0.8)	24 (1.6)	28 (2.0)	<0.001
Yes							
Sex with a CSW							
No	15,464 (92.6)	5,768 (87.4)	2,686 (92.3)	4,131 (97.4)	1,480 (97.6)	1,399 (98.0)	
Yes	1,232 (7.4)	835 (12.6)	224 (7.7)	109 (2.6)	36 (2.4)	28 (2.0)	<0.001
Exchanged sex for money or drugs							
No	16,319 (97.7)	6,425 (97.3)	2,791 (95.9)	4,201 (99.1)	1,503 (99.1)	1,399 (98.0)	
Yes	377 (2.3)	178 (2.7)	119 (4.1)	39 (0.9)	13 (0.9)	28 (2.0)	<0.001
Number of sex partners							
0	773 (4.6)	150 (2.3)	53 (1.8)	390 (9.2)	115 (7.6)	65 (4.5)	
1	4,409 (26.4)	1,527 (23.3)	535 (18.4)	1,372 (32.4)	454 (30.0)	521 (36.5)	
2–4	5,234 (31.4)	2,640 (40.0)	1,008 (34.6)	756 (17.8)	311 (20.5)	519 (36.4)	
5–9	1,810 (10.8)	1,039 (15.8)	359 (12.3)	205 (4.8)	74 (4.9)	133 (9.3)	
≥10	1,418 (8.5)	659 (10.0)	443 (15.2)	165 (3.9)	50 (3.3)	101 (7.1)	
Missing	3,052 (18.3)	588 (8.9)	512 (17.6)	1,352 (31.9)	512 (33.8)	88 (6.2)	<0.001
Condom use with current sex partner							
Always	2,293 (13.7)	588 (8.9)	719 (24.7)	496 (11.7)	179 (11.8)	311 (21.8)	
Sometimes	8,062 (48.3)	4,609 (69.8)	1,268 (43.6)	1,255 (29.6)	450 (29.7)	480 (33.6)	
Never	4,693 (28.1)	1,170 (17.7)	711 (24.4)	1,859 (43.9)	482 (31.8)	471 (33.0)	
Missing	1,648 (9.9)	236 (3.6)	212 (7.3)	630 (14.9)	405 (26.7)	165 (11.6)	<0.001

continued on p. 613

Table 2 (continued). Patient risks by site of HIV test, 2002–2004[a]

	Total	*UCC*	*Drop-In Center*	*Inpatient*	*PCC*	*ED*	
	N *(percent)*	N *(percent)*	N *(percent)*	N *(percent)*	N *(percent)*	N *(percent)*	P-*value*
Anal-insertive sex (men only)							
Did not have	6,359 (65.3)	3,401 (81.2)	585 (33.4)	1,375 (57.9)	631 (80.2)	367 (58.4)	
<6 weeks ago	467 (4.8)	173 (4.1)	186 (10.6)	56 (2.4)	38 (4.8)	14 (2.2)	
≥6 weeks ago	525 (5.4)	295 (7.0)	97 (5.5)	67 (2.8)	42 (5.3)	24 (3.8)	
Missing	2,383 (24.5)	322 (7.7)	885 (50.5)	876 (36.9)	76 (9.7)	224 (35.6)	<0.001
Anal-receptive sex							
Did not have	13,813 (82.7)	6,099 (92.4)	1,746 (60.0)	3,315 (78.2)	1,405 (92.7)	1,248 (87.5)	
<6 weeks ago	190 (1.1)	62 (0.9)	85 (2.9)	11 (0.3)	15 (1.0)	17 (1.2)	
≥6 weeks ago	195 (1.2)	86 (1.3)	57 (2.0)	17 (0.4)	9 (0.6)	26 (1.8)	
Missing	2,498 (15.0)	356 (5.4)	1,022 (35.1)	897 (21.2)	87 (5.7)	136 (9.5)	<0.001
Vaginal sex							
Did not have	999 (6.0)	267 (4.0)	244 (8.4)	348 (8.2)	83 (5.5)	57 (4.0)	
<6 weeks ago	9,437 (56.5)	4,328 (65.6)	1,563 (53.7)	1,984 (46.8)	879 (58.0)	683 (47.9)	
≥6 weeks ago	4,829 (28.9)	1,617 (24.5)	766 (26.3)	1,511 (35.6)	418 (27.6)	517 (36.2)	
Missing	1,431 (8.6)	391 (5.9)	337 (11.6)	397 (9.4)	136 (9.0)	170 (11.9)	<0.001

[a]All risks refer to past three years except where noted: receptive and insertive anal sex, vaginal sex, and condom use (no reference period specified). Analysis limited to patients with valid HIV test results available, and those who did not test HIV-positive subsequent to a previous HIV-positive test result.

HIV = human immunodeficiency virus

UCC = Urgent Care Center

PCC = primary care clinic

ED = Emergency Department

IDU = injection drug user

STD = sexually transmitted disease

CSW = commercial sex worker

Programmatic costs

The programmatic cost per HIV test and per previously unreported HIV infection was lowest in the UCC ($30.29 per HIV test and $1,980.00 per HIV-positive test result). The UCC had the highest HIV prevalence detected (1.5%), and also the greatest number of patients tested per eight-hour shift (6.3). While the Drop-In Center had the same HIV prevalence as the UCC, it had the lowest number of patients tested per eight-hour shift (2.8), contributing to its higher cost per HIV test ($68.74) and cost per HIV-positive test result ($4,652.00). Although the ED had a higher number of HIV tests per eight-hour shift (5.5) than inpatient (4.0), PCC (2.8), and the Drop-In Center (4.3), it had the lowest prevalence (0.6%) and highest cost per HIV-positive test result ($9,724) (Table 4).

DISCUSSION

We found a high prevalence (>1%) of HIV among patients accepting voluntary testing in four out of five sites within our medical center. Self-reported risk varied considerably by site. The costs per HIV test conducted and HIV-positive test result were lowest in the UCC. The cost per HIV test was highest at the Drop-In Center, which had the lowest number of tests per eight-hour shift. The cost per HIV-positive test result was highest in the ED, which had the lowest HIV prevalence. These results are supported by a recent analysis comparing the cost-effectiveness of targeted testing to CDC's recommendations for mass testing.[18] Targeted testing was more cost-effective through greater number of infections detected and more transmissions prevented.[18] This analysis also took into account the trade-off in saved time and resources for opt-out testing compared with the potential reductions in transmissions through targeted risk assessment and counseling, and found that targeted counseling and testing were more cost-effective, even when the HIV prevalence was as low as 0.3%.[18]

Our comparison of HIV testing in multiple sites within a single medical center allows for a discussion of certain advantages and disadvantages of targeted and universal HIV testing. The detection of all cases of HIV infection would require routine HIV testing of all patients, as CDC recommends. To facilitate this,

Table 3. Number and percent HIV-positive test results by site and patient characteristics

	HIV-positive N	*Percent HIV-positive (95% CI)*[a]	*Unadjusted odds ratio (95% CI)*
Site			
Emergency Department	8	0.56 (0.24, 1.10)	Ref.
Primary care clinic	18	1.19 (0.71, 1.87)	2.13 (0.93, 4.92)
Inpatient floor	59	1.39 (1.06, 1.79)	2.50 (1.19, 5.15)
Drop-In Center	43	1.48 (1.07, 1.99)	2.66 (1.55, 5.67)
Urgent Care Center	101	1.53 (1.25, 1.86)	2.76 (1.34, 5.67)
Gender			
Female	78	1.15 (0.91, 1.43)	Ref.
Male	145	1.49 (1.26, 1.75)	1.30 (0.98, 1.71)
Age (in years)			
15–24	14	0.48 (0.26, 0.80)	Ref.
25–39	106	1.52 (1.25, 1.84)	3.22 (1.84, 5.64)
40–54	93	1.85 (1.50, 2.26)	3.39 (2.24, 6.91)
≥55	8	0.57 (0.25, 1.12)	1.19 (0.50, 2.85)
Race			
White	24	0.63 (0.40, 0.93)	Ref.
Black	100	1.56 (1.27, 1.89)	2.51 (1.61, 3.93)
Hispanic	39	1.22 (0.87, 1.67)	1.97 (1.18, 3.28)
Haitian	30	2.36 (1.60, 3.35)	3.84 (2.24, 6.60)
Cape Verdean	10	2.09 (1.01, 3.81)	3.40 (1.16, 7.15)
Asian	1	0.31 (0.00, 1.72)	0.50 (0.07, 3.68)
Other, unknown	25	2.11 (1.37, 3.10)	3.43 (1.96, 6.02)
Health insurance status			
Insured	167	1.24 (1.06, 1.44)	Ref.
Not insured	45	2.05 (1.50, 2.73)	1.66 (1.19, 2.32)
Education			
≥High school/GED	126	1.13 (0.95, 1.35)	Ref.
<High school	69	2.00 (1.56, 2.53)	1.78 (1.33, 2.02)
Declined, missing	34	1.59 (1.10, 2.21)	1.41 (0.96, 2.07)
Referral			
Self	135	1.22 (1.02, 1.44)	Ref.
Physician	72	1.83 (1.44, 2.31)	1.51 (1.13, 2.02)
Other	22	1.29 (0.81, 1.94)	1.05 (0.67, 1.66)
Sex partner gender			
Opposite gender only	178	1.24 (1.07, 1.44)	Ref.
Same gender or both	34	2.66 (1.85, 3.69)	2.17 (1.49, 3.15)
Injection drug use			
No	208	1.35 (1.18, 1.55)	Ref.
Yes	21	1.59 (0.98, 2.41)	1.17 (0.75, 1.85)
History of STD diagnosis			
No	171	1.25 (1.07, 1.45)	Ref.
Yes	58	1.93 (1.47, 2.49)	1.56 (1.15, 2.11)
Sex with an HIV-positive person			
No	205	1.27 (1.11, 1.46)	Ref.
Yes	24	4.03 (2.60, 5.93)	3.25 (2.11, 5.01)
Number of sex partners			
0	11	1.42 (0.71, 2.53)	Ref.
1	50	1.13 (0.83, 1.49)	0.79 (0.41, 1.53)
2–4	77	1.47 (1.16, 1.83)	1.03 (0.55, 1.95)
5–9	28	1.55 (1.03, 2.23)	1.09 (0.54, 2.20)
≥10	23	1.62 (1.03, 2.42)	1.14 (0.55, 2.36)
Missing	40	1.31 (0.94, 1.78)	0.92 (0.47, 1.80)
Condom use with current sex partner			
Always	30	1.31 (0.88, 1.86)	Ref.
Sometimes	120	1.49 (1.24, 1.78)	1.14 (0.76, 1.70)
Never	46	0.98 (0.72, 1.31)	0.75 (0.47, 1.19)
Missing	33	2.00 (1.38, 2.80)	1.54 (0.94, 2.54)

continued on p. 615

Table 3 (continued). Number and percent HIV-positive test results by site and patient characteristics

	HIV-positive N	*Percent HIV-positive (95% CI)*[a]	*Unadjusted odds ratio (95% CI)*
Number of sex partners			
0	11	1.42 (0.71, 2.53)	Ref.
1	50	1.13 (0.83, 1.49)	0.79 (0.41, 1.53)
2–4	77	1.47 (1.16, 1.83)	1.03 (0.55, 1.95)
5–9	28	1.55 (1.03, 2.23)	1.09 (0.54, 2.20)
≥10	23	1.62 (1.03, 2.42)	1.14 (0.55, 2.36)
Missing	40	1.31 (0.94, 1.78)	0.92 (0.47, 1.80)
Condom use with current sex partner			
Always	30	1.31 (0.88, 1.86)	Ref.
Sometimes	120	1.49 (1.24, 1.78)	1.14 (0.76, 1.70)
Never	46	0.98 (0.72, 1.31)	0.75 (0.47, 1.19)
Missing	33	2.00 (1.38, 2.80)	1.54 (0.94, 2.54)
Anal-insertive sex (men only)			
Did not have	75	1.18 (0.93, 1.48)	Ref.
<6 weeks ago	10	2.14 (1.03, 3.90)	1.83 (0.94, 3.57)
≥6 weeks ago	20	3.81 (2.34, 5.82)	3.32 (2.01, 5.48)
Missing	40	1.68 (1.20, 2.28)	1.43 (0.97, 2.11)
Anal-receptive sex			
Did not have	161	1.17 (0.99, 1.36)	Ref.
<6 weeks ago	5	2.63 (0.86, 6.03)	2.29 (0.93, 5.65)
≥6 weeks ago	14	7.18 (3.98, 11.8)	6.56 (3.73, 11.5)
Missing	49	1.96 (1.45, 2.59)	1.70 (1.23, 2.34)
Vaginal sex			
Did not have	31	3.10 (2.12, 4.38)	Ref.
<6 weeks ago	99	1.05 (0.85, 1.28)	0.33 (0.22, 0.50)
≥6 weeks ago	74	1.53 (1.21, 1.92)	0.49 (0.32, 0.74)
Missing	25	1.75 (1.13, 2.57)	0.56 (0.33, 0.95)
Year of HIV test			
2002	98	2.27 (1.84, 2.76)	Ref.
2003	61	1.27 (0.97, 1.62)	0.55 (0.40, 0.76)
2004	70	0.93 (0.72, 1.17)	0.40 (0.30, 0.55)

[a]95% CIs for HIV prevalence are binomial exact CIs.

HIV = human immunodeficiency virus

CI = confidence interval

Ref. = reference group

GED = graduated equivalent degree

STD = sexually transmitted disease

CDC has suggested eliminating pretest counseling and separate written consent.[3] While these process modifications should mitigate several logistical barriers, it remains to be seen whether they will be sufficient to permit routine HIV testing in high-volume settings.

In our medical center, inpatient and outpatient volume far exceed counselor resources (e.g., UCC volume is approximately 80 patients daily, with one to two counselors offering testing in this area). We propose that in an opt-out approach with no risk assessment or pretest counseling, person-time would still be required, at minimum (*1*) to inform patients they will be tested for HIV, (*2*) for specimen collection and processing, (*3*) for delivering test results, (*4*) for documentation, and (*5*) for linkage to care. Even if this process took only a few minutes per person, a clinician-delivered program (i.e., nurses or physicians) is likely to be time- and cost-prohibitive. Thus, in our medical center, while we detected an HIV prevalence higher than 1% in all the testing sites except for the ED, it is not possible to offer testing to all patients due to the large volume of patients cared for and limited current resources.

Opt-out HIV screening for all patients in all clinical sites has not been conducted or evaluated in many clinical sites, including our own. A comprehensive programmatic evaluation would need to measure at each clinical site: numbers of patients declining opt-out testing, reasons for refusing, patient preferences for pretest counseling, counselor time involved in

specimen collection and processing, documentation, delivering results, and patient outcomes (follow-up with linkage to care). This information would quantify, and give descriptive value to, the effectiveness of CDC recommendations for HIV testing in medical settings. Additionally, as Girardi et al. report, while opt-out pretest counseling increased HIV testing volume in their medical center, the number of new infections detected remained constant compared with the period of targeted testing before, thus suggesting that opt-out testing may facilitate HIV testing among low-risk patients.[19]

Limitations

Results of the analysis were limited by data collected through the MDPH HIV Counseling and Testing Data Collection form. This analysis does not produce a true measure of HIV prevalence, because data represent the number of tests conducted rather than the number of individuals, and therefore are likely to include some patients who were tested more than once. There is evidence to suggest that patients who undergo repeat HIV testing are at higher risk for HIV infection.[20,21] In this case, our data would overestimate the prevalence of high-risk behaviors. We excluded from analysis individuals who reported their previous HIV test results as positive, but there may have been patients with previous HIV-positive test results who did not report it. Use of various HIV diagnostic tests may have introduced bias into the HIV prevalence detected. Demographic and behavioral risk information of people who were not offered testing (selection bias) or who declined testing (volunteer bias) may have differed from those who accepted HIV testing, which would affect our observed prevalence and associations. Given the mixture of confidential and anonymous testing, we cannot calculate the proportion of patients who entered into HIV primary care, an important aspect in evaluating program success.

Our cost analysis is limited because these data were collected for programmatic purposes. As such, we did not have data points that would have allowed a more detailed cost analysis. We estimated programmatic costs using a standardized FTEE salary for HIV counselors and did not account for material costs. Also, HIV testing for patients aged 15 to 54 in the ED was coupled with STD testing for patients aged 15 to 29. Seven hundred fifty-four (53%) ED patients were offered STD testing, and 546 (38%) were tested. Our simplified cost analysis does not take into account the extra time spent offering and testing for STDs. The person-time this took away from conducting additional HIV tests led to an overestimation of the cost per HIV test. However, even if twice as many HIV-positive test results had been detected, HIV testing in the ED would still have the greatest cost per HIV-positive test result ($4,862). We do not have measures of the personnel time spent in follow-up activities for HIV-positive people, nor were we able to calculate the cost per HIV-positive person entered into care. The current analysis highlights the

Table 4. Personnel costs and cost per HIV test and HIV-positive test result by site[a]

	Urgent Care Center	*Drop-In Center*	*Inpatient*	*Primary care clinic*	*Emergency Department*
Dates of observation	January 2002 to December 2004	January 2002 to December 2004	January 2002 to December 2004	January 2002 to December 2004	November 2003 to May 2004
Total number of patients tested for HIV	6,603	2,910	4,240	1,516	1,427
Total number of patients with HIV-positive results	101	43	59	18	8
Number of full-time employees	1.5	1.5	1.5	0.5	3.0
Total number of eight-hour shifts	1,057	1,057	1,057	352	259
Number of tests per eight-hour shift	6.2	2.8	4.0	4.3	5.5
Personnel costs	$200,025	$200,025	$200,025	$66,675	$77,788
Personnel costs per HIV test	$30.29	$68.74	$47.18	$43.98	$54.51
Personnel costs per HIV-positive test result	$1,980	$4,652	$3,390	$3,704	$9,724

[a]The number of eight-hour shifts was estimated for all sites except the Emergency Department, where they were documented.

HIV = human immunodeficiency virus

need for more detailed cost analysis, such as that conducted by Holtgrave et al.[18]

Our analysis found that the yield and cost of HIV testing and patient risk varied by clinical site. These results suggest that testing procedures and results from one clinical site might not necessarily be extrapolated to other clinical sites, even within a single medical center. Findings from our site also might not necessarily be extrapolated to other medical centers and geographic locales. Implementing CDC recommendations for routine HIV testing in clinical sites should take into consideration local procedures and resources.

CONCLUSIONS

Despite thousands of HIV tests having been conducted at this urban medical center over the past few years, and a significant number of HIV-positive individuals having been identified, the numbers of patients tested represent a minority (<10%) of patients evaluated clinically in these sites. Thus, it is likely that there were patients at risk for HIV who were not offered testing due to personnel constraints. Our results indicate HIV testing should be focused on patients who are uninsured, are younger than age 55, are males who report same-sex intercourse, and patients who report a history of STD or sex with an HIV-positive person. Clinicians play a significant role in identifying people at risk for HIV, as demonstrated by the approximately one-third of all HIV-positive people who were identified by physician referral.

While it is important to offer HIV information and testing to all patients, if the main goal is to identify HIV-infected individuals, our results indicate that pragmatically, HIV testing should remain targeted until a feasible and resource-efficient system that preserves linkage to care is in place for testing all patients for HIV infection.

Human immunodeficiency virus (HIV) counseling and testing in the primary care clinic, Urgent Care Center, Drop-In Center, and the inpatient medical service was funded through a contract with the Massachusetts Department of Public Health, HIV/AIDS Bureau (49304350099, PI Skolnik). Data collection in the Emergency Department was funded by a cooperative agreement from the Centers for Disease Control and Prevention (CDC), Division of HIV/AIDS Prevention (RR18/CCR120999, PI Mehta).

The findings and conclusions in this report are those of the authors and do not necessarily represent the views of CDC. The funding sources had no role in the analysis and interpretation of these data or the preparation, review, or approval of this article.

For Emergency Department data collection, CDC project officers were involved in the design and conduct of the study. Data were collected and analyzed while Dr. Mehta was affiliated with the Department of Emergency Medicine at Boston University School of Medicine.

REFERENCES

1. Glynn M, Rhodes P. Estimated HIV prevalence in the United States at the end of 2003. National HIV Prevention Conference; 2005 Jun 12–15; Atlanta. Abstract 595.
2. Centers for Disease Control and Prevention (US). HIV/AIDS surveillance report, 2003 (Vol. 15). Atlanta: Department of Health and Human Services, CDC (US); 2004.
3. Advancing HIV prevention: new strategies for a changing epidemic—United States, 2003. MMWR Morb Mortal Wkly Rep 2003; 52(15):329-32.
4. Revised recommendations for HIV testing of adults, adolescents, and pregnant women in health-care settings. MMWR Recomm Rep 2006;55(RR-14):1-17.
5. Walensky RP, Losina E, Malatesta L, Barton GE, O'Connor CA, Skolnik PR, et al. Effective HIV case identification through routine HIV screening at urgent care centers in Massachusetts. Am J Public Health 2005;95:71-3.
6. Walensky RP, Losina E, Steger-Craven KA, Freedberg KA. Identifying undiagnosed human immunodeficiency virus: the yield of routine, voluntary inpatient testing. Arch Intern Med 2002;162:887-92.
7. Paltiel AD, Weinstein MC, Kimmel AD, Seage GR 3rd, Losina E, Zhang H, et al. Expanded screening for HIV in the United States—an analysis of cost-effectiveness. N Engl J Med 2005;352:586-95.
8. Greenwald JL, Rich CA, Bessega S, Posner MA, Maeda JL, Skolnik PR. Evaluation of the Centers for Disease Control and Prevention's recommendations regarding routine testing for human immunodeficiency virus by an inpatient service: who are we missing? Mayo Clin Proc 2006;81:452-8.
9. Greenwald JL. Routine rapid HIV testing in hospitals: another opportunity for hospitalists to improve care. J Hosp Med 2006;1: 106-12.
10. Routinely recommended HIV testing at an urban urgent care clinic—Atlanta, Georgia, 2000. MMWR Morb Mortal Wkly Rep 2001;50(25):538-41.
11. Voluntary HIV testing as part of routine medical care—Massachusetts, 2002. MMWR Morb Mortal Wkly Rep 2004;53(24):523-6.
12. Kelen GD, Shahan JB, Quinn TC. Emergency department-based HIV screening and counseling: experience with rapid and standard serologic testing. Ann Emerg Med 1999;33:147-55.
13. Alpert PL, Shuter J, DeShaw MG, Webber MP, Klein RS. Factors associated with unrecognized HIV-1 infection in an inner-city emergency department. Ann Emerg Med 1996;28:159-64.
14. Glick NR, Silva A, Zun L, Whitman S. HIV testing in a resource-poor urban emergency department. AIDS Educ Prev 2004;16:126-36.
15. Pincus JM, Crosby SS, Losina E, King ER, LaBelle C, Freedberg KA. Acute human immunodeficiency virus infection in patients presenting to an urban urgent care center. Clin Infect Dis 2003; 37:1699-704.
16. Mehta SD, Hall J, Lyss SB, Skolnik PR, Pealer LN, Kharasch S. Adult and pediatric emergency department STD and HIV screening: programmatic overview and outcomes. Acad Emerg Med 2007; 14:250-8.
17. StataCorp. Stata statistical software: Release 9.2. College Station (TX): StataCorp.; 2005.
18. Holtgrave DR. Costs and consequences of the US Centers for Disease Control and Prevention's recommendations for opt-out HIV testing. PLoS Med 2007;4:e194.
19. Girardi E, Puro V, De Carli G, Orchi N, Ippolito G. Consent policies and rates of HIV testing. JAMA 2007;297:2478-9.
20. Fernyak SE, Page-Shafer K, Kellogg TA, McFarland W, Katz MH. Risk behaviors and HIV incidence among repeat testers at publicly funded HIV testing sites in San Francisco. J Acquir Immune Defic Syndr 2002;31:63-70.
21. MacKellar DA, Valleroy LA, Secura GM, Bartholow BN, McFarland W, Shehan D, et al. Repeat HIV testing, risk behaviors, and HIV seroconversion among young men who have sex with men: a call to monitor and improve the practice of prevention. J Acquir Immune Defic Syndr 2002;29:76-85.

Unhealthy Drinking Patterns and Receipt of Preventive Medical Services by Older Adults

Elizabeth L. Merrick, PhD, MSW[1], Dominic Hodgkin, PhD[1], Deborah W. Garnick, ScD[1], Constance M. Horgan, ScD[1], Lee Panas, MS[1], Marian Ryan, RRT, MPH, CHES[1], Richard Saitz, MD, MPH[2,3,4], and Frederic C. Blow, PhD[5,6]

[1]Institute for Behavioral Health, Heller School for Social Policy and Management, Brandeis University, Waltham, MA, USA; [2]Boston University School of Medicine, Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Boston Medical Center, Boston, MA, USA; [3]Youth Alcohol Prevention Center, Boston University School of Public Health, Boston, MA, USA; [4]Department of Epidemiology, Boston University School of Public Health, Boston, MA, USA; [5]Department of Psychiatry, University of Michigan, Ann Arbor, MI, USA; [6]Department of Veterans Affairs, Health Services Research and Development, Serious Mental Illness Treatment Research and Evaluation Center, Ann Arbor, MI, USA.

BACKGROUND: Preventive service use among older adults is suboptimal. Unhealthy drinking may constitute a risk factor for failure to receive these services.

OBJECTIVES: To determine the relationship between unhealthy drinking and receipt of recommended preventive services among elderly Medicare beneficiaries, applying the framework of current alcohol consumption guidelines.

DESIGN/METHODS: The data source is the nationally representative 2003 Medicare Current Beneficiary Survey. The sample included community-dwelling, fee-for-service Medicare beneficiaries 65 years and older (N = 10,523). Based on self-reported drinking, respondents were categorized as nondrinkers, within-guidelines drinkers, exceeding monthly but not daily limits, or heavy episodic drinkers. Using survey and claims data, influenza vaccination, pneumonia vaccination, glaucoma screening, and mammogram receipt were determined. Bivariate and logistic regression analyses were conducted.

RESULTS: Overall, 70.3% received flu vaccination and 49% received glaucoma screening during the year, 66.8% received pneumonia vaccination, and 56.2% of women received a mammogram over 2 years. In logistic regression, heavy episodic drinking was associated with lower likelihood of service receipt compared to drinking within guidelines: flu vaccination (OR 0.75, CI 0.59–0.96), glaucoma screening (OR 0.74, CI 0.58–0.95), and pneumonia vaccination (OR 0.75, CI 0.59–0.96). Nondrinkers when compared with those reporting drinking within guidelines were less likely to receive a mammogram (OR 0.83, CI 0.69–1.00).

CONCLUSIONS: Heavy episodic drinking is associated with lower likelihood of receiving several preventive services. Practitioners should be encouraged to screen all elders regarding alcohol intake and in addition to appropriate intervention, consider elders reporting heavy episodic drinking at higher risk for non-receipt of preventive services.

KEY WORDS: unhealthy drinking; Medicare beneficiaries; preventive services; older adults.
J Gen Intern Med 23(11):1741–8
DOI: 10.1007/s11606-008-0753-3

Received December 11, 2007
Revised June 11, 2008
Accepted July 1, 2008
Published online August 21, 2008

INTRODUCTION

Preventive medical services for older adults can improve their health and quality of life, through preventing life-threatening disease (such as with influenza and pneumonia vaccination) and early detection (through cancer screening).[1] Despite recommendations for preventive services by organizations such as the U.S. Preventive Services Task Force (USPSTF), elders' receipt of services is suboptimal. For example, one study found that although 91% of female Medicare beneficiaries aged 65 years and older receive at least one preventive service, only 10% receive all that are appropriate.[2] Another found that one- to two-thirds of elderly Medicare beneficiaries did not receive most recommended preventive services.[3] Underuse of preventive services among elders represents a growing public health problem, given the U.S. Census estimate that by 2030 one in five Americans will be 65 years of age or over.[4] The problem is increasingly recognized, as evidenced by Medicare initiatives to cover more preventive services[5] and to include preventive services in "pay-for-performance" approaches to physician payment.[6]

Numerous studies have confirmed low preventive services use among Medicare beneficiaries and identified correlates.[3,7–12] Lower service use is associated with Medicaid coverage, lack of supplemental insurance, less than high school education, lower income, smoking, being non-married, female gender, black race, Hispanic ethnicity, and inadequate health literacy; results for health status and behavioral health factors have been mixed.

The impact of behavioral health problems, and unhealthy alcohol use in particular, on preventive service use among elders is an important research area. Unhealthy alcohol use encompasses risky use, problem drinking, and alcohol disorders, including abuse and dependence.[13] About 9% of

community-dwelling elderly Medicare beneficiaries report drinking that exceeds recommended guidelines.[14] The National Institute on Alcohol Abuse and Alcoholism (NIAAA) and American Geriatrics Society (AGS) define risky drinking for those 65 years and older as more than seven drinks per week, or more than three drinks on any single day.[15,16] Exceeding these limits is associated with interpersonal and functioning problems for elders,[17] who have higher sensitivity and impaired ability to metabolize alcohol.[13]

Excessive alcohol use could affect preventive services use in several ways. Persons with unhealthy alcohol use could be physically or cognitively impaired in ways that reduce their ability to access appropriate services. Alternatively, excessive drinking may indicate generalized self-neglect in terms of health.[18,19] Providers might also treat this population differently.[20,21]

There is limited empirical work on the impact of alcohol use on older adults' preventive service use. A study of the elderly Medicare population found that very heavy drinkers – those drinking at least four drinks per night on eight or more nights per month – used fewer preventive services overall, though results for specific services were not reported.[3] Another study of adults 55 and older found that harmful drinkers were less likely than social drinkers to receive a pneumococcal vaccination.[10] A study of women aged 50 and older found that those who consumed alcoholic beverages had higher mammography rates than non-drinkers, but did not examine effects of heavier drinking specifically.[22] Research including, but not limited to the elderly has found a negative relationship between preventive service receipt and substance use.[23,24] The current study fills this gap in the literature by applying the framework of current alcohol consumption guidelines to the use of widely recommended preventive services, using nationally representative data.

METHODS

Data and Sample

The primary data source is the 2003 Access to Care file of the Medicare Current Beneficiary Survey (MCBS), conducted continuously from 1991.[25] The sample is selected using a stratified, multistage probability sample design to represent the Medicare population nationally. MCBS sample weights are provided to achieve nationally representative estimates. The survey is based on in-person interviews administered three times per year. Content includes sociodemographics, health and functional status, and health-care utilization. The 2003 MCBS included items regarding alcohol consumption and preventive care services.[26] Subjects' Medicare claims are also provided and were linked to survey data for this analysis. The 2003 MCBS Access to Care sample consisted of 16,003 beneficiaries. This study included community-dwelling persons 65 years of age or older. Health maintenance organization (HMO) enrollees were excluded because their claims were not available. After excluding 3,520 subjects under 65 years of age or institutionalized, another 1,890 subjects enrolled in HMOs, and 59 subjects missing alcohol data, this study included 10,523 persons representing a weighted N of 26,617,034. Sample sizes varied by sub-analysis due to item-missing data (<2.5 to 8%) or selecting women for the mammogram analysis who were also in the 2002 MCBS so utilization could be observed over 2 years.

Measures

Alcohol Consumption Variables. The 2003 MCBS included three alcohol consumption items. Quantity and frequency were ascertained by asking, "Please think about a typical month in the past year. On how many days did you drink any type of alcoholic beverage? On those days that you drank alcohol, how many drinks did you have?" Heavy episodic drinking was assessed by asking, "Please think about a typical month in the past year. On how many days did you have 4 or more drinks in a single day?" Alcoholic beverages were described as including "liquor such as whiskey or gin, mixed drinks, wine, beer, and any other type of alcoholic beverage."

To assess unhealthy drinking in terms of consuming risky amounts of alcohol (regardless of whether alcohol problems or disorders were present), we defined alcohol measures reflecting two parameters of the NIAAA[16] and AGS guidelines.[15] First, to be consistent with the weekly guideline, we defined "exceeding monthly limits" as more than 30 drinks per typical month. Twenty-six respondents reporting 31 drinks per month whose responses were clearly based on a 31-day month were also coded as negative, since the items did not specify standardized number of days per month. Second, we constructed a "heavy episodic drinking" variable, indicating whether an individual reported four or more drinks in any single day during a typical month in the past year, according to either drinking quantity item.

We categorized respondents into four mutually exclusive categories: non-drinkers; within-guidelines drinkers (not exceeding the monthly limit or the three-drink single-day limit); drinkers who exceeded the monthly limit, but not the single-day limit; and heavy episodic drinkers who exceeded the single-day drinking limit, with or without exceeding the monthly limit. For descriptive purposes, we also calculated continuous measures of drinking quantity and frequency.

Covariates. Covariates were selected that were identified previously to affect health-care utilization. Sociodemographic variables included gender, race, Hispanic ethnicity, annual household income, age, education, marital status, and residence in a metropolitan area. Living arrangement was not included due to high correlation with marital status.

We controlled for health status by utilizing DxCG (diagnostic cost group) risk-adjustment software[27] that uses sex, age, and diagnosis codes from claims to construct a continuous measure of relative risk of health-care resource use. Compared to other illness burden indices or scales, the DxCg score contains higher specificity related to the individual's clinical profile in projecting future health-care costs and estimating an individual's care management needs.[28,29] Thus, it may be used as a proxy for health status in that higher DxCG risk scores denote higher health-care resource use risk and presumably poorer health[30] A value of 1 indicates the individual's predicted cost equals the population average for all persons with Medicare claims; higher values indicate higher-than-average predicted costs. For bivariate analyses, we created categories: no claims or claims not indicative of significant health risk (DxCG score

<0.1); claims indicative of lower-than-average health risk (0.1 ≤ DxCG score ≤ 1); claims indicative of higher-than-average health risk (DxCG score >1). For logistic regression models, we used the continuous measure; increasing scores indicate higher risk of health-care resource use (poorer health status).

We controlled for functional status utilizing a modified Katz Index of Independence in Activities of Daily Living (ADL)[31] variable constructed from survey data. Respondents were asked whether they had trouble or needed assistance with six ADLs: bathing, dressing, transferring, toileting, continence, or feeding. If no difficulty was indicated, that activity received a score of 1. The resulting variable reflects a 7-position scale (0–6) of the number of independent ADLs.

Two mental health variables were used. First, a self-reported depression variable was created. Respondents were asked, "In the past 12 months, how much of the time did you feel sad, blue, or depressed?" (all, most, some, little, or none of the time), and "In the past 12 months, have you had 2 weeks or more when you lost interest or pleasure in things that you usually cared about or enjoyed?" (yes, no). Respondents who answered "all" or "most of the time" to the first question and/or "yes" to the second question were categorized as having self-reported depression. This covariate was constructed to approximate the modified PHQ-2 validated for older adults.[32] Second, a dichotomous variable indicating mental health diagnosis was constructed based on a single ICD-9 coded claim with a mental health diagnosis in the inpatient setting or two such claims in the outpatient setting during the year.

Other dichotomous variables included current smoking, having a usual source of care, private supplemental insurance, and Medicaid coverage (all survey-based).

PREVENTIVE SERVICE MEASURES

Dichotomous variables were created for four preventive services that are widely recommended for elders with no definitive age cut-off and were feasible to measure with this dataset.

Influenza vaccination. The USPSTF[33] and Centers for Disease Control (CDC)[34] recommend annual vaccination for adults 50 years of age and older. A positive response was based on the question, "Did you get a flu shot any time during the period from September (previous year) through December (previous year)?"

Glaucoma screening. While the USPSTF found insufficient evidence to recommend routine glaucoma screening for persons over 65 years of age,[33] it was recommended by other national experts [The American Academy of Ophthalmologists,[35] National Committee for Quality Assurance (NCQA),[36] and the Veterans Administration[35]]. We include it as an example of a widely, but not universally recommended preventive care measure for a major health problem in this population. We constructed a claims-based measure to determine glaucoma screening receipt during the year, following NCQA's Healthcare Effectiveness Data and Information Set (HEDIS) specifications.[36] HEDIS measures are widely endorsed performance measures accompanied by technical specifications for calculating from administrative data.

Pneumonia vaccination. The USPSTF and CDC recommend that all adults 65 years of age and older receive a one-time pneumonia vaccination.[33,34] A positive response was based on the question, "Have you ever had a shot for pneumonia?"

Mammogram. The USPSTF recommends that women over 40 years of age receive a mammogram every 1–2 years.[33] There is no definitive age cut-off, but rather decision-making for those over age 70 should be guided by a woman's life expectancy given her health status. Prior research found that survey data tend to overestimate mammogram receipt compared to claims data.[37] Therefore, we used a claims-based measure to determine mammogram receipt during a 2-year period according to HEDIS specifications, without an age ceiling.[36] The variable was positive if a mammogram claim was identified in either 2002 or 2003 among women present in both years of the MCBS.

Statistical Analysis

Results presented here are weighted estimates that represent the continuously enrolled, community-dwelling, non-HMO, elderly Medicare population. Chi-square tests were used to assess bivariate differences by drinking category; chi-square statistics were corrected for the survey design and converted to F-statistics. We conducted logistic regression analyses to model receipt of each service as a function of alcohol consumption and covariates.

Due to the complex sampling design, using procedures that assumed equal probability of selection would likely lead to underestimating standard errors.[38,39] The SVY:LOGIT procedure of the statistical package STATA version 9.0 was used to more accurately determine the statistical significance of the observed differences (STATA Corporation, College Station, TX).

RESULTS

The weighted sample reflected a predominantly white (87.7%), non-Hispanic (93.7%) population living in a metropolitan area (73.7%) (Table 1). Most (56.5%) were female. Almost one-third (30.6%) were aged 65–70, 45.9% were 71–80, and 23.5% were 81 years or older.

Two-thirds (65.5%) of the sample reported drinking no alcohol during a typical month in the past year (Table 1). One-quarter (25.4%) reported drinking within guidelines, 3.8% exceeded the monthly limit only, and 5.4% reported heavy episodic drinking (exceeding the three-drink daily limit, with or without exceeding the monthly limit). Among persons exceeding the monthly limit only, drinks per month ranged from 32 to 93 with a mean of 60.4 (SD 12.1) and median of 60. Among heavy episodic drinkers, the range was 0.5 to 625 (a value of 0.5 was assigned to the response of "less than one drink" and based on the frequency-quantity series, not the separate binge item); the mean was 63.1 (SD 69.3) and the median was 50. The mean number of heavy drinking days was 7.6 (SD 9.5, median 3.0).

In bivariate analyses, heavy episodic drinkers were significantly less likely ($p < .001$) to receive flu vaccination, glaucoma screening, or pneumonia vaccination than all other drinking categories (Table 2). Mammogram receipt showed a different

Table 1. Sample Description

Weighted percent	Total	Non-drinker	Drinks within guidelines	Drinks over monthly limit	Heavy episodic drinking	P value*
Unweighted n	10,523	7,114	2,539	379	491	
Weighted N	26,617,034	17,422,361	6,760,633	1,004,758	1,429,282	
Total		65.5	25.4	3.8	5.4	
Female	56.5	63.3	50.5	31.3	20.5	<.001
Age						<.001
65 to 70 years	30.6	27.7	33.7	34.4	48.7	
71 to 80 years	45.9	45.4	47.7	47.5	41.8	
≥81 years	23.5	26.9	18.6	18.1	9.5	
Hispanic	6.3	7.5	3.6	†	7.9	<.001
Race						<.001
White	87.7	84.6	93.4	97.7	90.5	
Black	8.2	10.3	4.0	†	7.4	
Asian	1.8	2.2	1.2	†	†	
Other	2.3	2.9	1.4	†	†	
Education						<.001
< H.S. diploma	29.1	35.9	15.3	14.4	29.1	
H.S. graduate	29.9	31.4	27.7	25.5	26.5	
Some college/voc/tech	23.9	21.5	29.6	22.9	27.5	
College degree	17.1	11.3	27.5	37.2	24.1	
Annual income						<.001
<$25,000	57.3	66.8	39.5	27.4	44.4	
$25,000 to $40,000	22.2	20.1	27.2	23.7	24.7	
>$40,000	20.5	13.2	33.3	48.9	31.0	
Marital status						<.001
Married	55.8	50.8	65.1	68.2	64.3	
Widowed	32.5	37.5	25.2	21.4	14.5	
Divorced, separated, single	11.7	11.7	9.7	10.4	21.2	
Metropolitan area	73.7	69.4	82.4	79.8	79.3	<.001
Region						<.001
Northeast	18.9	17.9	21.7	21.4	16.3	
South	37.8	41.4	30.7	31.0	32.0	
Midwest	25.9	25.2	27.9	23.3	26.6	
West	15.9	13.5	19.3	24.2	22.1	
Other	1.5	1.9	0.4	†	3.0	
Relative health risk						<.001
No risk indication	31.0	28.6	33.6	36.0	44.6	
< Average risk	42.7	41.7	46.5	43.4	37.6	
> Average risk	26.3	29.8	20.0	20.7	17.7	
Self-reported depression	13.2	16.0	7.6	6.3	10.9	<.001
Mental health dx	13.9	16.1	9.3	8.5	11.6	<.001
Current smoker	11.2	10.4	10.1	11.6	25.7	<.001
Functional status						<.001
Independence in 6 ADLs	60.6	56.8	66.4	74.1	69.8	
Independence in 5 ADLs	23.8	24.0	24.3	20.1	21.0	
Independence in 0 to 4 ADLs	15.7	19.2	9.3	5.8	9.2	
Usual source of care	95.3	95.5	95.7	94.9	90.1	<.001
Private insurance	73.6	69.6	82.8	84.6	70.4	<.001
Medicaid coverage	12.2	16.2	4.5	2.0	7.6	<.001

*Chi-square tests, corrected for survey design; weighted sample used
†Not shown due to unweighted cell n<11

pattern, being significantly lower for nondrinkers (52.8%) than for within-guidelines drinkers (66.3%) ($p < .001$).

In the logistic regression model predicting flu vaccination, heavy episodic drinking (as compared to within-guideline drinking) was significantly associated with lower odds of vaccination (OR 0.75, CI 0.59–0.96, p=0.02) and glaucoma screening (OR 0.74, CI 0.58–0.95, p=0.02) (Table 3). Heavy episodic drinking was also associated with lower odds of pneumonia vaccination (OR 0.75, CI 0.59–0.96, p=0.02) (Table 4). Nondrinkers were less likely to receive a mammogram (OR 0.83, CI 0.69.- 1.00, p=0.05) compared to within-guideline drinkers; heavy episodic drinking and drinking over the monthly limit were not significant (Table 4).

Sensitivity analyses using survey data, or survey plus claims data, failed to find an association between nondrinkers and mammogram receipt (data not shown). Analyses employing interaction terms for gender by drinking category into the models for the other three services were not significant in the flu vaccination and glaucoma screening models. For pneumonia vaccination, only the interaction variable for "female by exceeding monthly guidelines" was significant and was positively associated with vaccination receipt (p<.01). The heavy

Table 2. Utilization of Preventive Services By Drinking Pattern

	Flu vaccination	Glaucoma screening	Pneumonia vaccination	Mammogram
Weighted percent*				
Unweighted n	10,469	10,523	10,387	4,051
Weighted N	26,484,972	26,617,034	26,235,718	9,809,367
Overall percent	70.3	49.0	66.8	56.2
Drinking category	†	†	†	†
Non-drinker	69.5	49.0	66.0	52.8
Drinks within guidelines	73.7	51.8	70.2	66.3
Drinks over monthly limit	75.1	52.0	71.8	65.9
Heavy episodic	60.1	34.3	57.7	63.2
Gender		†		N/A
Female	70.4	53.6	67.0	56.2
Male	70.1	43.0	66.6	N/A
Age	†	†	†	†
65 to 70 years	61.4	35.8	55.1	66.3
71 to 80 years	73.1	53.0	71.2	59.6
≥81 years	76.2	58.4	73.4	41.5
Ethnicity	†	†	†	†
Hispanic	52.6	37.8	49.5	42.9
Non-Hispanic	71.4	49.8	68.0	57.2
Race	†	†	†	†
White	72.0	50.4	69.1	58.0
Black	55.2	37.2	52.1	42.9
Other	62.8	42.6	49.1	46.9
Education	†	†	†	†
< H.S. diploma	62.5	42.9	60.0	44.6
H.S. graduate	72.0	50.2	67.1	57.8
Some college/voc/tech	72.8	50.6	72.5	62.2
College degree	76.9	55.2	70.3	68.2
Annual income	†	†	†	†
<$25,000	66.9	47.7	63.6	49.9
$25,000 to $40,000	71.6	50.1	71.0	65.0
>$40,000	78.7	52.3	72.4	71.5
Marital status	†	†	†	†
Married	72.3	48.4	68.2	66.9
Widowed	70.4	53.8	67.6	49.9
Divorced, separated, single	60.2	38.8	58.1	45.1
Metropolitan area				
In metropolitan area	70.1	49.4	68.0	56.9
Not in metro area	70.6	47.9	66.4	53.9
Region	†	‡	†	‡
Northeast	72.0	52.0	64.7	53.7
South	69.3	50.0	67.6	56.3
Midwest	72.5	47.9	68.0	59.9
West	70.9	46.6	68.0	54.9
Other	27.8	30.3	43.3	37.2
Relative health risk	†	†	†	†
No risk indication	59.5	33.5	56.5	48.5
< Average risk	74.0	56.2	70.7	62.2
> Average risk	76.8	55.6	72.8	54.1
Self-reported depression				†
Depression	68.4	47.2	66.7	48.7
No depression	70.5	49.3	67.3	57.5
Mental health dx	†		‡	‡
MH diagnosis	75.2	51.0	70.0	52.1
No MH diagnosis	69.5	48.7	66.3	57.0
Current smoker	†	†	†	†
Current smoker	59.3	34.7	59.7	39.4
Not current smoker	71.6	50.8	67.7	57.8
Functional status	‡		†	†
Independence in 6 ADLs	69.0	49.6	64.6	58.5
Independence in 5 ADLs	72.3	51.2	69.9	60.9
Independence in 0 to 4 ADLs	71.9	49.6	71.0	43.0
Usual source of care	†	†	†	†
Had usual source	72.1	50.1	68.4	57.5
No usual source	34.6	26.2	34.3	26.9
Private insurance	†	†	†	†
Private insurance	74.0	53.3	70.4	61.5
No private insurance	59.6	37.1	56.8	41.1
Medicaid coverage	†	†	†	†
Medicaid	61.8	44.8	53.5	41.6
No Medicaid	71.4	49.6	68.7	58.8

*Weighted sample used; significance based on chi-square tests, corrected for survey design
† <.01
‡<.05

Table 3. Logistic Regression Results*: Flu Vaccination and Glaucoma Screening

	Odds ratio (95% confidence interval)	
	Flu vaccination	Glaucoma screening
Drinking category		
Drinks within guidelines	Reference	Reference
Non-drinker	0.99 (0.87 – 1.13)	0.95 (0.85 – 1.07)
Drinks over monthly limit only	0.99 (0.75 – 1.31)	1.12 (0.87 – 1.45)
Heavy episodic drinking	0.75 (0.59 – 0.96) ‡	0.74 (0.58 – 0.95) ‡
Female	0.98 (0.88 – 1.09)	1.50 (1.37 – 1.66) †
Age		
65 to 70 years	Reference	Reference
71 to 80 years	1.64 (1.46 – 1.86) †	1.91 (1.71 – 2.14) †
≥81	1.87 (1.61 – 2.18) †	2.31 (2.03 – 2.63) †
Hispanic	0.78 (0.60 – 1.00)	0.88 (0.70 – 1.10)
Race		
White	Reference	Reference
Black	0.62 (0.51- 0.75) †	0.69 (0.57 – 0.84) †
Other	0.87 (0.65 – 1.16)	1.01 (0.78 – 1.30)
Education		
< HS diploma	Reference	Reference
High school graduate	1.32 (1.15 – 1.51) †	1.22 (1.10 – 1.36) †
Some college/voc/tech	1.40 (1.23 – 1.59) †	1.29 (1.13 – 1.47) †
College degree	1.58 (1.33 – 1.87) †	1.53 (1.31 – 1.80) †
Annual income		
<$25,000	Reference	Reference
Income $25,000 to $40,000	1.04 (0.90 – 1.20)	1.34 (0.92 – 1.15)
Income >$40,000	1.40 (1.19 – 1.65) †	1.06 (0.92 – 1.22)
Marital status		
Married	Reference	Reference
Divorced, separated, or single	0.75 (0.64 – 0.86) †	0.81 (0.69 – 0.95) †
Widowed	0.90 (0.79 – 1.04)	0.99 (0.89 – 1.10)
Metropolitan area	0.92 (0.81 – 1.05)	1.01 (0.89 – 1.14)
Region		
South	Reference	Reference
Northeast	1.16 (0.98 – 1.36)	1.01 (0.88 – 1.16)
Midwest	1.10 (0.96 – 1.27)	0.86 (0.76 – 0.98) ‡
West	1.10 (0.91 – 1.34)	0.85 (0.76 – 0.95) †
Other	0.27 (0.18 – 0.41) †	0.58 (0.23 – 1.46)
Relative health risk	1.10 (1.05 – 1.15) †	1.07 (1.03 – 1.11) †
Self-reported depression	0.95 (0.81 – 1.12)	0.96 (0.84 – 1.10)
Current smoker	0.78 (0.67 – 0.91) †	0.68 (0.58 – 0.80) †
Functional status		
Independence in 6 ADLs	Reference	Reference
Independence in 5 ADLs	1.10 (0.96 – 1.26)	1.02 (0.90 – 1.15)
Independence in 0 to 4 ADLs	1.08 (0.93 – 1.25)	0.88 (0.77 – 1.00)
Usual source of care	3.97 (3.11 – 5.07) †	2.29 (1.78 – 2.96) †
Mental health diagnosis	1.23 (1.07 – 1.42) †	0.99 (0.86 – 1.16)
Private supplemental insurance	1.51 (1.31 – 1.73) †	1.86 (1.64 – 2.11) †
Medicaid coverage	1.28 (1.05 – 1.56) ‡	1.62 (1.41 – 1.92) †

*Weighted sample used
†<.01
‡<.05

episodic drinking variable retained its main effect at p<.05 in all three models.

DISCUSSION

Heavy episodic drinking, as defined by current guidelines, was associated with lower likelihood of receiving a preventive service in three out of four types of preventive care services we examined. Exceeding monthly, but not single-day limits was not associated with less preventive service receipt. There are several possible explanations. Persons with heavy episodic drinking are more likely to have diagnosable alcohol use disorders.[40] Their lower likelihood of preventive services may be part of a constellation of behaviors reflecting self-neglect and/or impaired judgment. They may be difficult to engage, or their drinking may present competing demands that result in less clinical time for encouraging preventive care. Previous research found that heavier drinking was associated with fewer physician visits, offering less opportunity for preventive

Table 4. Logistic Regression Results*: Pneumonia Vaccination and Mammogram

	Odds ratio (95% confidence interval)	
	Pneumonia vaccination	Mammogram
Drinking category		
Drinks within guidelines	Reference	Reference
Non-drinker	0.94 (0.83 – 1.07)	0.83 (0.69 – 1.00) ‡
Drinks over monthly limit only	1.05 (0.80 – 1.38)	1.03 (0.59 - 1.79)
Heavy episodic drinking	0.75 (0.59 – 0.96) ‡	0.86 (0.50 – 1.47)
Female	1.00 (0.91 – 1.09)	
Age		
65 to 70 years	Reference	Reference
71 to 80 years	1.97 (1.76 – 2.22) †	0.75 (0.63 – 0.91)
≥81	2.16 (1.88 – 2.47) †	0.39 (0.32 – 0.47) †
Hispanic	0.72 (0.55 – 0.94) ‡	0.94 (0.67 – 1.31) †
Race		
White	Reference	Reference
Black	0.62 (0.51 – 0.74) †	0.80 (0.63 – 1.03)
Other	0.63 (0.48 – 0.82) †	0.73 (0.48 – 1.10)
Education		
< HS diploma	Reference	Reference
High school graduate	1.17 (1.02 – 1.34) ‡	1.21 (1.00 – 1.47) ‡
Some college/voc/tech	1.51 (1.32 1.72) †	1.40 (1.12 – 1.74) †
College degree	1.23 (1.04 – 1.46) ‡	1.62 (1.22 – 2.16) †
Annual household income		
<$25,000	Reference	Reference
Income $25,000 to $40,000	1.22 (1.05 – 1.41) †	1.10 (0.90 – 1.34)
Income >$40,000	1.28 (1.11 – 1.48) †	1.31 (0.98 – 1.74)
Marital status		
Married	Reference	Reference
Divorced, separated, or single	0.88 (0.75 – 1.03)	0.49 (0.38 – 0.64) †
Widowed	0.95 (0.84 – 1.08)	0.72 (0.61 – 0.85) †
Metropolitan area	0.91 (0.77 – 1.07)	1.14 (0.95 – 1.36)
Region		
South	Reference	Reference
Northeast	0.88 (0.74 – 1.04)	0.97 (0.77 – 1.23)
Midwest	0.95 (0.81 – 1.11)	1.11 (0.91 – 1.35)
West	1.06 (0.88 – 1.27)	0.92 (0.74 – 1.13)
Other	0.68 (0.40 – 1.16)	0.69 (0.35 – 1.33)
Relative health risk	1.10 (1.05 – 1.14) †	1.01 (0.96 – 1.06)
Self-reported depression	1.05 (0.89 – 1.24)	0.90 (0.73 – 1.11)
Current smoker	0.94 (0.79 – 1.12)	0.42 (0.31 – 0.57) †
Functional status		
Independence in 6 ADLs	Reference	Reference
Independence in 5 ADLs	1.22 (1.09 – 1.37) †	1.11 (0.93 – 1.32)
Independence in 0 to 4 ADLs	1.28 (1.10 – 1.48) †	0.68 (0.56 – 0.83) †
Usual source of care	3.53 (2.86 – 4.35) †	3.95 (2.73 – 5.71) †
Mental health diagnosis	1.05 (0.90 – 1.22)	1.00 (0.80 – 1.24)
Private supplemental insurance	1.25 (1.09 – 1.44) †	1.83 (1.45 – 2.31) †
Medicaid coverage	0.85 (0.71 – 1.02)	1.34 (1.03 – 1.74) ‡

*Weighted sample used
†<.01
‡<.05

services.[41] Alternatively, heavy episodic drinkers may have providers who focus less on preventive care.[42]

The lack of association between heavy episodic drinking and mammogram receipt is unique among the four services. Only nondrinkers were significantly less likely to receive the service than within-guidelines drinkers. The finding is somewhat puzzling, though consistent with some prior research.[22] Although moderate (and heavy) drinking is a known risk factor for breast cancer, it seems unlikely that drinkers sought mammography because of this risk as it is not well publicized.[43] Most analyses employing interaction terms did not indicate gender differences in the relationship between alcohol consumption and preventive service receipt. Certainly fewer women report heavy drinking than men.[44,45] Further research is needed to confirm the role of older women's drinking in preventive services use, and if confirmed, to understand why this might vary by service.

The effect of heavy episodic drinking was similar for services that are consistently recommended and that Medicare covers universally (influenza and pneumonia vaccinations), as well as for glaucoma screening, for which recommendations are mixed and Medicare covers only for high-risk groups. Thus, factors other than professional consensus and extent of coverage may drive the relationship between heavy episodic drinking and preventive service receipt. The relatively low preventive service use overall suggests that implementing multiple strategies to improve service delivery, as Medicare initiatives are aiming to do, is warranted.

The lack of significance found between over-monthly-limit drinking and receipt of preventive services may reflect a population at risk for chronic problems, but without current impairment affecting preventive services use. In contrast, heavy episodic drinking is more likely associated with at least acute cognitive impairment, which can lead to social disorganization and general self-neglect. This would be a fruitful area for further research.

This study has several limitations. The study's cross-sectional design does not permit determination of causality. MCBS data are not ideally suited to precise dose-response analyses, which we therefore did not conduct. Other measures also carry some imprecision, including mental health variables: claims underestimate prevalence of mental health disorders, and self-reported depression is not synonymous with clinical disorder. The DxCG risk score is only a proxy for health status, but sensitivity analyses using self-reported health status did not change key findings (data not shown). Our analyses included glaucoma screening, a service not universally recommended. However, a similar pattern was observed across several measures. Finally, it is worth noting that study findings neither validate nor cast doubt on the alcohol guidelines, which were developed with a range of outcomes in mind, not health-care utilization.

The study goal was to examine the relationship between alcohol guideline adherence and preventive services receipt. Results suggest that elders with heavy episodic drinking are at risk for failure to receive certain recommended preventive services. Health-care providers and others working with older adults should be alert to the broad range of problems associated with unhealthy drinking and be encouraged to screen proactively all elders regarding alcohol consumption. Investigation of underlying causal mechanisms is needed. Nonetheless, currently recommended screening for unhealthy alcohol use could also identify those at risk for not receiving indicated preventive services, and interventions directed at lowering consumption might also improve preventive service use.

Acknowledgments: *This study was funded by the National Institute on Alcohol Abuse and Alcoholism grant no. 5R21AA015746. Preliminary findings from the study were presented at the annual conference of the American Public Health Association, 7 November 2007. The authors thank Grant Ritter, Ph.D., for statistical consultation and Michele Hutcheon for manuscript preparation.*

Conflict of Interest: *None disclosed.*

Corresponding Author: *Elizabeth L. Merrick, PhD, MSW; Institute for Behavioral Health, Heller School for Social Policy and Management, Brandeis University, Mail stop 035, 415 South St., Waltham, MA 02454, USA (e-mail: merrick@brandeis.edu).*

REFERENCES

1. **Goldberg TH, Chavin SI.** Preventive medicine and screening in older adults. J Am Geriatr Soc. 1997;45(3):344–54.
2. U.S. Government Accounting Office. Medicare preventive services, most beneficiaries receive some but not all recommended services. GAO Report. 2004;1–12. Available at: http://www.gao.gov/new.items/d041004t.pdf. Accessed June 5, 2008.
3. **Ozminkowski RJ, Goetzel RZ, Shechter D, Stapleton DC, Baser O, Lapin P.** Predictors of preventive service use among Medicare beneficiaries. Health Care Financ Rev. 2006;27(3):5–23.
4. US Census Bureau News. Census Bureau projects tripling of Hispanic and Asian populations in 50 years; non-Hispanic whites may drop to half of total population. Available at: http://www.census.gov/Press-Release/www/releases/archives/population/001720.html. Accessed June 5, 2008.
5. Centers for Medicare and Medicaid Services. Medicare program; revisions to payment policies under the physician fee schedule for calendar year 2005. Fed Regist. 2004;69(219):66235–66915. Available at: http://www.access.gpo.gov/su_docs/fedreg/a041115c.html. Accessed June 5, 2008.
6. Centers for Medicare and Medicaid Services. Medicare "pay for performance (P4P)" initiatives. Available at: http://www.cms.hhs.gov/apps/media/press/release.asp?Counter=1343. Accessed June 5, 2008.
7. **Kosiak B, Sangl J, Correa-de-Araujo R.** Quality of health care for older women: what do we know? Women's Health Issues. 2006;16:89–99.
8. **Pham HH, Schrag D, Hargraves JL, Bach PB.** Delivery of preventive services to older adults by primary care physicians. JAMA. 2005;294(4):473–81.
9. **Asch SM, Sloss EM, Hogan C, Brook RH, Kravitz RL.** Measuring underuse of necessary care among elderly Medicare beneficiaries using inpatient and outpatient claims. JAMA. 2000;284(18):2325–33.
10. **Moore AA, Morgenstern H, Harawa NT, Fielding JE, Higa J, Beck JC.** Are older hazardous and harmful drinkers less likely to participate in health-related behaviors and practices as compared with nonhazardous drinkers? J Am Geriatr Soc. 2001;49(4):421–30.
11. **Scott TL, Gazmararian JA, Williams MV, Baker DW.** Health literacy and preventive health care use among Medicare enrollees in a managed care organization. Med Care. 2002;40(5):395–404.
12. **Klabunde CN, Meissner HI, Wooten KG, Breen N, Singleton JA.** Comparing colorectal cancer screening and immunization status in older Americans. Am J Prev Med. 2007;33(1):1–8.
13. **Saitz R.** Medical and Surgical Complications of Addiction. In: Graham AST, Mayo-Smith M, Ries R, Wilford BB, ed. Princ. of Addic Med. 3rd ed.: Chevy Chase, MD: ASAM. 2003:1027–52.
14. **Merrick EL, Horgan CM, Hodgkin D, Garnick DW, Houghton SF, Panas L, Saitz R.** Unhealthy drinking patterns among older adults: prevalence and associated characteristics. JAGS. 2008;56:214–23.
15. American Geriatrics Society Clinical Practice Committee. Clinical guidelines for alcohol use disorders in older adults. Available at: http://www.americangeriatrics.org/products/positionpapers/alcohol.shtml. Accessed June 5, 2008.

16. National Institute on Alcohol Abuse and Alcoholism. Helping patients who drink too much. A Clinician's Guide. 2005 Edition. Available at: http://pubs.niaaa.nih.gov/publications/Practitioner/CliniciansGuide2005/guide.pdf. Accessed June 5, 2008.
17. **Moos RH, Brennan PL, Schutte KK, Moos BS.** High-risk alcohol consumption and late-life alcohol use problems. Am J Public Health. 2004;94(11):1985–91.
18. **Branch LG, Rabiner DJ.** Rediscovering the patient's role in receiving health promotion services. Med Care. 2000;38(1):70–7.
19. **Rabiner DJ, Branch LG, Sullivan RJ Jr.** Patient factors related to the odds of receiving prevention services in Veterans Health Administration medical centers. Am J Manag Care. 1999;5(9):1153–60.
20. **Rush B, Ellis K, Crowe T, Powell L.** How general practitioners view alcohol use - Clearing up the confusion. Can Fam Physician. 1994;40:1570–9.
21. **Warburg MM, Cleary PD, Rohman M, Barnes HN, Aronson M, Delbanco TL.** Residents' attitudes, knowledge, and behavior regarding diagnosis and treatment of alcoholism. J Med Educ. 1987;62(6):497–503.
22. **Fredman L, Sexton M, Cui Y, et al.** Cigarette smoking, alcohol consumption, and screening mammography among women ages 50 and older. Prev Med. 1999;28:407–17.
23. **Cryer PC, Jenkins LM, Cook AC, et al.** The use of acute and preventative medical services by a general population: relationship to alcohol consumption. Addiction. 1999;94(10):1523–32.
24. **Druss BG, Rosenheck RA, Desai MM, Perlin JB.** Quality of preventive medical care for patients with mental disorders. Med Care. 2002;40(2):129–36.
25. Centers for Medicare and Medicaid Services. Medicare current beneficiary survey overview. Available at: http://www.cms.hhs.gov/MCBS/Downloads/CNP_2003_appendixA.pdf. Accessed June 5, 2008.
26. Centers for Medicare and Medicaid Services. MCBS main study - round 37 - fall supplement 2003 community component HF. Health Status and Functioning; 2003. Available at: http://www.cms.hhs.gov/MCBS/Downloads/2003_CBQ_hs.pdf Accessed June 5, 2008.
27. DxCG. Products and services, RiskSmart Stand Alone. Available at: http://www.dxcg.com/products-services/risksmart-stand-alone.asp. Accessed June 5, 2008.
28. **Zhao Y, Ellis RP, Ash AS, et al.** Measuring population health risks using inpatient diagnoses and outpatient pharmacy data. Health Serv Res. 2001;36(6 Pt 2):180–93.
29. **Zhao Y, Ash A, Ellis RP, Slaughter JP.** Disease burden profiles: an emerging tool for managing managed care. Health Care Manag Sci. 2002;5:211–9.
30. **Wang MC, Rosen AK, Kazis L, Loveland S, Anderson J, Berlowitz D.** Correlation of risk adjustment measures based on diagnoses and patient self-reported health status. Health Serv Outcomes Res Method. 2000;1(3–4):351–65.
31. The Hartford Institute for Geriatric Nursing, Best Practices in Nursing Care to Older Adults. Katz Index of Independence in Activities of Daily Living (ADL). Available at: http://www.hartfordign.org/publications/trythis/issue02.pdf. Accessed June 5, 2008.
32. **Li C, Friedman B, Conwell Y, Fiscella K.** Validity of the patient health questionnaire 2 (PHQ-2) in identifying major depression in older people. J Am Geriatr Soc. 2007;55:596–602.
33. US Preventive Services Task Force. Recommendations. Available at: http://www.ahrq.gov/clinic/USPSTFix.htm#Recommendations. Accessed June 5, 2008.
34. Centers for Disease Control. Recommended adult immunization schedule. Available at: http://www.cdc.gov/vaccines/recs/schedules/downloads/adult/06–07/adult-schedule-11x17-bw.pdf. Accessed June 5, 2008.
35. **Goldberg LD.** Clinical guidelines for the treatment of glaucoma. Manag Care. 2002;11(S):16–24.
36. National Committee for Quality Assurance. HEDIS 2005 Volume 2, Technical Manual. Washington DC, 2004.
37. **Randolph WM, Mahnken JD, Goodwin JS, Freeman JL.** Using Medicare data to estimate the prevalence of breast cancer screening in older women: comparison of different methods to identify screening mammograms. Health Serv Res. 2002;37(6):1643–57.
38. **Cohen SB.** An evaluation of alternative PC-based software packages developed for the analysis of complex survey data. Am Stat. 1997;51(3):285–92.
39. **Lemeshow S, Letenneur L, Dartigues J, Lafont S, Orgogozo J, Commenges D.** Illustration of analysis taking into account complex survey considerations: the association between wine consumption and dementia in the PAQUID study. Am J Epidemiol. 1998;148(3):298–306.
40. National Institute on Alcohol Abuse and Alcoholism. National epidemiological survey on alcohol and related conditions. Alcohol Alert. Available at: http://pubs.niaaa.nih.gov/publications/AA70/AA70.htm. Accessed June 5, 2008.
41. **Rice C, Duncan D.** Alcohol use and reported physician visits in older adults. Prev Med. 1995;24:229–34.
42. **Heywood A, Sanson-Fisher R, Ring I, Mudge P.** Risk prevalence and screening for cancer by general practitioners. Prev Med. 1994;23:152–9.
43. National Cancer Institute. Estimating breast cancer risk: questions and answers. NCI Fact Sheets. Available at: http://www.cancer.gov/cancertopics/factsheet/estimating-breast-cancer-risk. Accessed June 5, 2008.
44. **Armstrong MA, Midanik LT, Klatsky AL.** Alcohol consumption and utilization of health services in a health maintenance organization. Med Care. 1998;36(11):1599–605.
45. **Midanik LT, Clark WB.** The demographic distribution of U.S. drinking patterns in 1990: description and trends from 1984. Am J Public Health. 1994;84(8):1218–22.

Atherosclerosis 203 (2009) 509–514

Contents lists available at ScienceDirect

Atherosclerosis

journal homepage: www.elsevier.com/locate/atherosclerosis

Cross-sectional relations of multiple inflammatory biomarkers to peripheral arterial disease: The Framingham Offspring Study☆

Joanne M. Murabito[a,b,*], Michelle J. Keyes[a,f], Chao-Yu Guo[g], John F. Keaney Jr.[h], Ramachandran S. Vasan[a,c,d,e,i], Ralph B. D'Agostino Sr.[a,f], Emelia J. Benjamin[a,c,d,e,i]

[a] National Heart, Lung and Blood Institute's Framingham Heart Study, Framingham, MA, United States
[b] Section of General Internal Medicine, Boston University School of Medicine, Boston, MA, United States
[c] Department of Cardiology, Boston University School of Medicine, Boston, MA, United States
[d] Department of Preventive Medicine, Boston University School of Medicine, Boston, MA, United States
[e] Whitaker Cardiovascular Institute, Boston University School of Medicine, Boston, MA, United States
[f] Statistics and Consulting Unit, Department of Mathematics and Statistics, Boston University, Boston, MA, United States
[g] Children's Hospital and Harvard Medical School, Boston, MA, United States
[h] University of Massachusetts Medical School, Worcester, MA, United States
[i] Epidemiology Section, Boston University School of Public Health, Boston University, Boston, MA, United States

ARTICLE INFO

Article history:
Received 7 March 2008
Received in revised form 19 June 2008
Accepted 27 June 2008
Available online 9 July 2008

Keywords:
Peripheral arterial disease
Ankle-brachial index
Interleukin-6
Tumor necrosis factor receptor 2

ABSTRACT

Background: Emerging evidence suggests that different inflammatory biomarkers operate through distinct biologic mechanisms. We hypothesized that the relation to peripheral arterial disease (PAD) varies for individual markers.

Methods: In a community-based sample we measured 12 biomarkers including *plasma* CD40 ligand, fibrinogen, lipoprotein-associated phospholipase-A2 mass and activity, osteoprotegerin, P-selectin, and tumor necrosis factor receptor 2 (TNFR2); and *serum* C-reactive protein, intracellular adhesion molecule-1, interleukin-6, monocyte chemoattractant protein-1, and myeloperoxidase in Framingham Offspring Study participants ($n = 2800$, 53% women, mean age 61 years). We examined the cross-sectional relation of the biomarker panel to PAD using (1) a global test of significance to determine whether at least one of 12 biomarkers was related to PAD using the TEST statement in the LOGISTIC procedure in SAS and (2) stepwise multivariable logistic regression with forward selection of markers with separate models for (1) ankle-brachial index (ABI) category (<0.9, 0.9–1.0, >1.0) and (2) presence of clinical PAD (intermittent claudication or lower extremity revascularization).

Results: The group of inflammatory biomarkers were significantly related to both ABI and clinical PAD ($p = 0.01$ and $p = 0.02$, respectively, multi-marker adjusted global significance test). Multivariable forward elimination regression retained interleukin-6 and TNFR2 as significantly associated with PAD. For one standard deviation change in interleukin-6 and TNFR2 concentrations, there was a 1.21 ($p = 0.005$) and 1.19 ($p = 0.009$) increased odds of a change in ABI level respectively. Similar results were observed for clinical PAD.

Conclusion: Interleukin-6 and TNFR2 were significantly associated with PAD independent of established risk factors and each other, suggesting that each marker represents a distinct biologic pathway.

☆ The Framingham Heart Study is supported by NIH/NHLBI contract N01-HC-25195. Additional funding for this work supported by NIH grants RO1 HL064753, HL076784, AG028321 (E.J.B.); HL 080124, HL 077477, HL71039, K24 HL 4334 (R.S.V.); tumor necrosis factor-alpha concentrations measured via American Diabetes Association Career Development Award and NCRR GCRC M01-RR-01066 James B. Meigs, MD, MPH. Lipoprotein-associated phospholipase-A2 Lp-PLA2 activity measurements were provided by GlaxoSmithKline at no cost to the FHS.

* Corresponding author at: The Framingham Heart Study, 73 Mount Wayte Avenue, Suite 2, Framingham, MA 01702-5827, United States. Tel.: +1 508 935 3461; fax: +1 508 626 1262.
E-mail address: murabito@bu.edu (J.M. Murabito).

0021-9150/$ – see front matter © 2008 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.atherosclerosis.2008.06.031

1. Introduction

Peripheral arterial disease (PAD) affects approximately five to eight million Americans [1] and is a powerful predictor of incident coronary heart disease, stroke, and mortality [2–4]. The ankle-brachial index (ABI), a subclinical measure of PAD, is considered a marker of generalized atherosclerosis. It is now well established that inflammation plays a central role in the pathogenesis of atherosclerosis and, further, that various inflammatory markers predict incident cardiovascular disease (CVD) events [5,6]. However, risk for CVD associated with inflammatory markers is often attenuated with adjustment for traditional risk factors.

The relation between inflammatory markers and PAD is not fully characterized. C-reactive protein predicts risk of symptomatic PAD [7,8], and also is associated with atherosclerosis in the aorta [9] and femoral artery [10], but the associations are attenuated after accounting for established risk factors. Adjusted inverse relations between the ABI and C-reactive protein have been reported in men [11], in ever smokers [12], and in persons with prevalent CVD [13]. Higher C-reactive protein concentrations are associated with progression of aortic, iliac, and lower extremity atherosclerosis [14], and in one small study both a low ABI and a high C-reactive protein identified persons at greatest risk for clinical events and death [15]. Reports of the relations between other inflammatory markers and PAD are limited, often focus on a single marker, or on small hospital-based or referral-based samples, and demonstrate conflicting results [10,16–18]. Emerging evidence suggests that different inflammatory markers operate through distinct biologic mechanisms, and thus the relative importance to the atherosclerotic process and PAD may differ for individual markers.

We examined the cross-sectional relations of a panel of 12 inflammatory biomarkers and PAD in a large community-based sample. We selected the inflammatory and oxidative stress markers to represent various stages and pathways in the inflammatory process, including chemokines (monocyte chemoattractant-1), cytokines (interleukin-6, tumor necrosis factor-alpha and tumor necrosis receptor-2; selectins [P-selectin and CF40 ligand] cell adhesion molecules [intercellular adhesion molecule-1]) acute phase reactants (C-reactive protein, fibrinogen), and an oxidative stress marker (myeloperoxidase). We hypothesized that different inflammatory markers represent distinct biologic pathways, and thus not all markers would be related to PAD. Moreover, we postulated that the strength of the relation would vary for different biomarkers operating through unique pathways. To our knowledge, no other study has evaluated the relation between multiple biomarkers from potentially diverse biologic pathways and PAD conjointly.

2. Methods

2.1. Study sample

The Framingham Offspring Study was initiated in 1971 when 5124 adult children (and offspring spouses) of the Original cohort were enrolled in the Framingham Heart Study. Offspring participants have been examined approximately every 4–8 years since the study's inception. Written informed consent was obtained at each examination and the Institutional Review Board of Boston University Medical Center approved the examination content.

Offspring participants who attended the seventh examination cycle (1998–2001) were eligible for this study. The examination included a standardized medical history and physical examination, electrocardiogram, noninvasive cardiovascular testing, and measurement of fasting lipids, glucose, and a panel of inflammatory biomarkers. Of the 3539 participants attending the examination, 205 were examined off-site and did not have ABI testing, 92 participants had incomplete ABI data, and 12 participants were excluded because of an ABI > 1.4. We further excluded participants missing all biomarker data ($n = 289$) and participants with incomplete risk factor data ($n = 141$). Thus, our study sample included 2800 Offspring participants with data available for all 12 biomarkers and complete risk factor data.

2.2. Measurement of ankle-brachial index

Ankle-brachial systolic blood pressure measurements were obtained using a standard protocol by trained technicians and the details previously published [19]. An 8 MHz Doppler pen probe and an ultrasonic Doppler flow detector (Parks Medical Electronics, Inc.) were used to measure the systolic blood pressure in each limb. All limb blood pressures were repeated, and if the initial and repeat blood pressures differed by more than 10 mmHg at any one site, a third measurement was obtained. Measurements were obtained from the dorsalis pedis artery only if the posterior tibial pulse could not be located by palpation or with Doppler probe. For this study, the ABI was defined as the ratio of the average systolic blood pressure in the ankle divided by the average systolic blood pressure in the higher arm. The lower ABI was used for analysis. Based on prior epidemiologic studies, we analyzed ABI < 0.9 as indicative of PAD.

2.3. Intermittent claudication and lower extremity revascularization

Intermittent claudication was assessed using a standardized physician-administered questionnaire that inquired about the presence of exertional calf discomfort related to walking uphill or walking rapidly and was relieved with rest. Two physicians independently interviewed all participants suspected to have intermittent claudication. An endpoint panel, comprised of three senior investigators, examined all medical evidence and made the final diagnosis of the presence of intermittent claudication. Participants were also queried about revascularization procedures including lower extremity bypass surgery and percutaneous transluminal angioplasty. The endpoint panel reviewed hospital records for all cardiovascular procedures.

2.4. Inflammatory biomarker measurement

At examination cycle seven, 12 biomarkers were measured including *plasma* CD40 ligand, fibrinogen, lipoprotein-associated phospholipase A2 mass and activity, osteoprotegerin, P-selectin, and tumor necrosis factor receptor 2 (TNFR2); and *serum* C-reactive protein, intracellular adhesion molecule-1, interleukin-6, monocyte chemoattractant protein-1, and myeloperoxidase. Specimens were collected from fasting participants and plasma and serum aliquots were stored at −80 °C until analysis. Biomarkers, except C-reactive protein, were measured in duplicate with commercially available ELISA kits from R&D Systems (intracellular adhesion molecule-1, interleukin-6, monocyte chemoattractant protein-1, P-selectin, TNFR2), Bender MedSystems (CD40 ligand), Diagnostica (fibrinogen), Oxis (myeloperoxidase) and ALPCO (osteoprotegerin). The Dade Behring BN100 nephelometer was used to measure high sensitivity C-reactive protein. Lipoprotein-associated phospholipase A2 activity was measured by GlaxoSmithKline, and mass was measured by DiaDexus. Details for assays have been previously published [20]. The intra-assay coefficients of variation for the biomarkers were as follows: CD40 ligand 4.4 ± 3.4%, fibrinogen 1.1 ± 1.1%, intracellular adhesion molecule-1 3.7 ± 2.4%, interleukin-6 3.1 ± 2.1%, lipoprotein-associated phospholipase A2 activity 7.0% (low) and 5.9% (high) and mass (based on 24% dupli-

cate readings) 6% (low) and 8% (high) concentrations, monocyte chemoattractant protein-1 3.8 ± 3.3%, myeloperoxidase 3.0 ± 2.5%, osteoprotegerin 3.7 ± 2.9%, P-selectin 3.0 ± 2.2%, TNFR2 2.2 ± 1.6%. The kappa statistic based on 146 C-reactive protein samples was 0.95. Additionally, plasma tumor necrosis factor alpha (R&D Systems, CV 7.6% low, 5.6% high control) and urinary isoprostanes, 8-Epi-$PGF_{2\alpha}$ (Cayman, Ann Arbor, MI; CV 9.1 ± 5.8%), indexed to urinary creatinine were measured on a subset of participants.

2.5. Clinical covariate assessment

Covariates were defined at the time of examination cycle seven. Medication use and current smoking within the year preceding the exam were self-reported. Resting blood pressure was measured twice by the examining physician. Hypertension was defined as an average blood pressure of systolic ≥140 or diastolic ≥90 mmHg or use of anti-hypertensive medication. Body mass index was calculated as weight in kilograms divided by the height in meters squared. Diabetes was defined by fasting blood glucose of ≥126 mg/dL, or use of insulin or oral hypoglycemic agents. CVD was defined as coronary heart disease, stroke or transient ischemic attack, and heart failure. An endpoint adjudication panel made the final diagnostic determination using previously reported criteria [21].

2.6. Statistical analysis

Sex-specific standardization of biomarkers was performed (i.e., within each sex, biomarkers were standardized to have a mean of 0 and a standard deviation of 1). Due to skewed distributions, biomarker concentrations were natural logarithmically transformed for analysis. Our primary analysis was the simultaneous consideration of multiple biomarkers (independent variables) in relation to PAD defined as two separate variables: (1) ABI category (ABI: <0.9, 0.9–1.0, >1.0) and (2) presence of clinically overt PAD defined as intermittent claudication or lower extremity revascularization. Separate logistic regression models were run for ABI category and presence of clinically overt PAD. First, we performed a global test of significance to determine whether at least one of 12 biomarkers was related to the PAD dependent variables using the TEST statement in the LOGISTIC procedure in SAS. The analysis was adjusted for age, sex, and the following 13 clinical covariates previously reported to be correlated with biomarkers and or PAD [19,22,23]: current cigarette smoking, number of pack-years of cigarette smoking, diabetes, fasting glucose, body mass index, waist circumference, total to HDL cholesterol ratio, fasting triglyceride, lipid lowering treatment, hypertension, aspirin use, prevalent CVD (myocardial infarction, coronary insufficiency, angina pectoris, stroke, or transient ischemic attack), and use of hormone replacement therapy. Second, we conducted a stepwise multivariable ordinal logistic regression with PAD as the dependent variable, with forward selection of biomarkers using a $p < 0.05$ adjusting for age, sex, and forcing the 13 clinical covariates into the model. For biomarkers identified to be related to PAD in the second step of the analysis we calculated point estimates of the odds ratio (or i.e., the relative change in odds of PAD), with 95% confidence intervals, per standard deviation increase of the biomarker examined.

We conducted several secondary analyses. We examined effect modification by age (<60, ≥60 years) and sex for significant biomarker—PAD relations. We repeated the analysis in persons free of CVD. Finally because multiple reports have used different markers or sets of markers, we analyzed the multivariable-adjusted linear relations of each log-transformed marker (dependent variable), one marker at a time, to the independent PAD measures using PROC GLM in SAS. Tumor necrosis factor-alpha was measured on a subset of participants attending examination cycle seven ($n = 2129$) and was included in the secondary analysis. SAS version 8.1 was used to perform all analyses [24].

3. Results

3.1. Participant characteristics and biomarker concentrations

Clinical characteristics of the study sample by presence of clinically overt PAD, and by ABI category are shown in Table 1. Participants with PAD, defined by symptoms or an ABI < 0.9, were older than participants without PAD. The untransformed median for the 12 biomarkers and tumor necrosis factor-alpha (available on a subset) by ABI level are shown in Table 2. A graded increase in marker concentrations across decreasing ABI levels was present for all markers except lipoprotein-associated phospholipase A2 mass and activity whereas an inverse relation was seen for CD40 ligand.

3.2. Global relations of multiple biomarkers and measures of PAD

The inflammatory markers as a group were significantly related to both ABI and clinically detected PAD ($p = 0.01$ and $p = 0.02$ respectively for the multi-marker adjusted global test of significance) as shown in Table 3. The forward elimination regression retained interleukin-6 and TNFR2 in the final models as significantly associated with ABI and with intermittent claudication or lower extremity revascularization. The odds of a one category reduction in ABI

Table 1
Clinical characteristics of the study sample

Variable mean (S.D.) or %	Intermittent claudication		Ankle-brachial index		
	Yes or vascular intervention, $N = 90$	No, $N = 2710$	<0.9, $N = 111$	0.9–1.0, $N = 225$	>1.0–1.4, $N = 2464$
Age, years	67 (9)	61 (9)	70 (8)	65 (10)	60 (9)
Women, %	43	54	50	71	52
Current smoking, %	27	13	30	26	12
Pack years among ever smokers, mean	63 (25)	43 (23)	65 (25)	47 (23)	41 (22)
Diabetes, %	29	12	26	20	11
Body mass index, kg/m^2	29.7 (6.1)	28.0 (5.1)	28.2 (5.5)	28.5 (6.6)	28.0 (4.9)
Total/HDL cholesterol ratio	4.5 (1.4)	4.0 (1.3)	4.4 (1.5)	4.0 (1.5)	4.1 (1.3)
Lipid lowering treatment, %	49	19	40	28	19
Hypertension, %	74	44	76	62	42
Hypertension treatment, %	68	32	66	49	31
Aspirin use[a], %	57	30	49	34	30
Prevalent cardiovascular disease[b], %	52	10	33	15	9
Hormone replacement among women, %	38	30	16	31	31

[a] Aspirin use is defined as three or more tablets per week.
[b] Cardiovascular disease did not include intermittent claudication.

Table 2
Unadjusted inflammatory marker data by ankle-brachial index level

Marker, units	Ankle-brachial index		
	<0.9, *N* = 111	0.9–1.0, *N* = 225	>1.0–1.4, *N* = 2464
	Untransformed marker concentrations, median (lower, upper quartile)		
CD40 ligand, ng/mL	0.78 (0.46, 2.31)	0.96 (0.53, 3.04)	1.27 (0.56, 4.07)
C-reactive protein, mg/L	3.77 (1.99, 8.89)	3.71 (1.53, 7.08)	2.02 (0.94, 4.76)
Fibrinogen, mg/dL	425 (384, 481)	388 (349, 446)	368 (326, 418)
Intercellular adhesion molecule-1, ng/mL	283 (243, 323)	255 (223, 292)	239 (209, 279)
Interleukin-6, pg/mL	4.82 (2.85, 7.84)	3.46 (2.21, 5.66)	2.58 (1.75, 4.09)
LpPLA2, mass, nmol/(mL min)	284 (231, 374)	293 (231, 367)	288 (229, 360)
LpPLA2, activity, ng/mL	143 (120, 173)	134 (116, 162)	141 (119, 165)
Monocyte chemoattractant protein-1, pg/mL	346 (283, 412)	326 (269, 409)	310 (252, 378)
Myeloperoxidase, ng/mL	43.4 (29.5, 61.7)	40.3 (26.9, 58.2)	39.9 (27.8, 59.6)
Osteoprotegerin, pmol/L	6.71 (5.35, 8.36)	5.85 (4.88, 7.10)	5.30 (4.39, 6.34)
P-selectin, pg/mL	39.7 (31.4, 54.0)	38.5 (31.4, 48.7)	36.0 (28.2, 45.1)
Tumor necrosis factor receptor 2, pg/mL	2407 (2024, 3187)	2154 (1718, 2856)	1945 (1642, 2340)
Tumor necrosis factor alpha[a], pg/mL	1.51 (1.15, 1.85)	1.37 (0.99, 1.92)	1.18 (0.92, 1.58)
Urine[a] 8-epi-$PGF_{2\alpha}$, ng/mmol	162 (103, 247)	151 (100, 234)	131 (88, 192)

[a] TNF-α data is available on a subset of 2129 participants, urine 8-epi-PGF_2 available on 2404 participants LpPLA2 = lipoprotein-associated phospholipase A2.

Table 3
Joint consideration of biomarkers in relation to the ankle-brachial index and clinical peripheral arterial disease

	Global P[a]	Stepwise selection biomarker[b]	Odds ratio (95% confidence interval)[c]	*p*-Value
ABI[d]	0.01	Interleukin-6	1.21 (1.06, 1.38)	0.005
		TNFR2	1.19 (1.05, 1.36)	0.009
Intermittent claudication or lower extremity revascularization	0.02	Interleukin-6	1.36 (1.06, 1.74)	0.02
		TNFR2	1.31 (1.04, 1.64)	0.02

[a] A simultaneous test of whether at least one of the 12 biomarkers were related to PAD (PAD is the dependent variable). Covariates in multivariable model include age, sex, current cigarette smoking, number of pack-years of cigarette smoking, diabetes, fasting glucose, body mass index, waist circumference, total to HDL cholesterol ratio, fasting triglyceride, lipid lowering treatment, hypertension, aspirin use (≥3 per week), prevalent cardiovascular disease (excluding intermittent claudication), and hormone replacement therapy use (women only).

[b] Individual biomarkers significantly related to PAD after forward stepwise selection (PAD is the dependent variable) are displayed.

[c] Point estimate indicates relative change in odds of PAD (ABI level or presence versus absence of intermittent claudication or lower extremity revascularization) per 1-standard deviation increment in log-marker (1-standard deviation increment is 0.71 for log Interleukin-6 and 0.30 for log TNFR2).

[d] The ABI was categorized as follows: <0.9, 0.9–1.0, >1.0.

level increased by 21 and 19% per a 1-standard deviation increase in interleukin-6 and TNFR2, respectively. Similar results were observed for intermittent claudication or revascularization.

3.3. Secondary analyses

In analyses of ABI, excluding participants with prevalent CVD ($n = 2496$), the global test examining whether the markers as a group were related to ABI was not significant ($p = 0.34$). The forward stepwise selection regression retained only interleukin-6 with a nearly identical point estimate (estimate 1.21, 95% confidence interval 1.05, 1.39, $p = 0.01$). The analysis of clinical PAD was not run in participants free of prevalent CVD due to small numbers ($n = 43$). No significant interactions were noted for sex and age with regard to the association between biomarkers and ABI.

In adjusted regression models examining each marker separately, C-reactive protein, interleukin-6, fibrinogen, tumor necrosis factor alpha, and TNFR2 were significantly inversely related to ABI level (p-values ranging from <0.0001 to 0.02). For each biomarker, we exponentiated the adjusted mean log-transformed biomarker and its 95% confidence interval to obtain its adjusted geometric mean and corresponding 95% confidence interval (Table 4). Similar markers were associated with clinical PAD (C-reactive protein, interleukin-6, and TNFR2; p-values ranging from <0.0001 to 0.01) with the following exceptions: fibrinogen and tumor necrosis factor alpha were not significantly associated (data not shown).

4. Discussion

4.1. Principal findings

In our cross-sectional community-based study, we examined the relations of a panel of 12 inflammatory biomarkers to PAD assessed by ABI, and by clinically defined intermittent claudication and/or lower extremity revascularization. Interleukin-6 and TNFR2 were significantly related to both measures of PAD. In secondary analyses, examining the relation of each marker separately to ABI, we observed additional significant inverse relations for C-reactive protein, fibrinogen, and tumor necrosis factor alpha after adjusting for known risk factors.

4.2. Interleukin-6 and PAD

Interleukin-6 is known to play a critical role in the inflammatory process with both pro-inflammatory and anti-inflammatory effects that include the stimulation of C-reactive protein, fibrinogen and other acute phase reactants and increased endothelial cell adhesiveness. In accordance with our results, in a small study of patients with intermittent claudication, interleukin-6 concentrations were higher in patients compared to healthy controls both at rest and after treadmill exercise ($p < 0.001$) suggesting that this marker is associated with peripheral atherosclerosis [25]. In a hospital-based investigation of the interleukin-6 G (-174) C geno-

Table 4
Secondary analyses: multivariable-adjusted regression of individual biomarkers on ankle-brachial index[a]

Biomarker	Geometric means and 95% confidence intervals[b]			
	Ankle-brachial index level			
	<0.9	0.9–1.0	>1.0–1.4	*p*-Value
C-reactive protein	2.91 (2.42, 3.51)	2.44 (2.15, 2.77)	2.19 (2.11, 2.28)	0.007
Interleukin-6	3.64 (3.21, 4.13)	3.09 (2.84, 3.37)	2.86 (2.78, 2.93)	0.0005
Fibrinogen	394 (381, 407)	372 (363, 381)	372 (369, 374)	0.005
Tumor necrosis factor alpha	1.43 (1.29, 1.58)	1.32 (1.24, 1.42)	1.25 (1.22, 1.27)	0.02
TNFR2	2258 (2142, 2379)	2087 (2013, 2164)	2009 (1987, 2030)	<0.0001

[a] Biomarkers with $p < 0.05$ displayed.
[b] For each biomarker (dependent variable), we exponentiated the adjusted mean log-transformed biomarker and its 95% confidence interval to obtain its adjusted geometric mean and corresponding 95% confidence interval. Covariates in multivariable model include covariates listed in Table 3 legend.

types, in patients with type II diabetes with and without PAD, the GG genotype and higher plasma concentrations of interleukin-6 and other inflammatory markers were more common in PAD patients [26]. The investigators of that report hypothesize that the GG genotype promotes PAD in patients with diabetes by inducing release of interleukin-6 which in turn results in increased concentrations of other biomarkers such as C-reactive protein. In the Edinburgh Artery Study, inflammatory marker concentrations, including interleukin-6, were significantly elevated at baseline in participants who developed symptomatic PAD during follow-up [27]. In that study, interleukin-6 was a predictor of incident PAD but the association was attenuated with adjustment for CVD risk factors.

Elevated concentrations of interleukin-6 have been noted in a community-based sample of older participants with a low ABI [28], a finding similar to our study. Furthermore, interleukin-6 was predictive of PAD progression defined by declining ABI over 12 years of follow-up even after adjusting for traditional risk factors and other inflammatory markers (C-reactive protein, intracellular adhesion molecule-1, vascular adhesion molecule-1, and E-selectin) [16], and hemostatic factors [17]. Moreover, interleukin-6 was the only inflammatory marker independently associated with ABI decline in persons free of baseline PAD. The independent predictive value of interleukin-6 in relation to PAD progression may reflect its role in both inflammatory and hemostatic processes. Additionally, interleukin-6 predicts the development of type II diabetes [29] and hypertension [30], both significant predictors of PAD. Finally, interleukin-6 predicts risk for incident CVD events [5] and persons with coronary disease have nearly a threefold risk of intermittent claudication. Hence, the association between interleukin-6 and PAD is likely mediated through a variety of complex interrelated biologic pathways and appears to extend to early peripheral atherosclerosis, atherosclerosis progression, and incident symptomatic disease.

4.3. TNFR2 and PAD

Tumor necrosis factor alpha is a pro-inflammatory cytokine that affects vascular tissues including endothelial cells. Tumor necrosis factor alpha exerts its biologic effects through two cell surface receptors, TNFR1 and TNFR2. However, the role of TNFR2 in the regulation of inflammatory responses in endothelial cells is unclear. In mice, the proatherogenic effect of tumor necrosis factor alpha was mediated primarily through TNFR2 [31]. Further, in mice endothelial TNFR2 is essential for tumor necrosis factor alpha induced leukocyte–endothelial-cell interaction which mediates several important steps of the inflammatory response including leukocyte rolling, adhesion, and transmigration [32]. A potential mechanism for TNFR2 mediated endothelial dysfunction is the down-regulation of lysyl oxidase, a key enzyme in extracellular matrix maturation. TNFR2 has been shown to be involved in lysyl oxidase down-regulation, which in turn is associated with endothelial dysfunction [33]. To our knowledge there is only one small study of patients with intermittent claudication and critical limb ischemia demonstrating elevated tumor necrosis factor receptor concentrations compared with controls [34].

4.4. Other markers and PAD

In the Physician's Health Study, C-reactive protein was the strongest nonlipid predictor of the development of symptomatic PAD [8]. In that report both C-reactive protein and fibrinogen improved risk prediction for PAD. However, the two markers were correlated and C-reactive protein was the stronger predictor of risk. The associations between C-reactive protein and fibrinogen and incident PAD were confirmed by the Edinburgh Artery Study and persisted after accounting for risk factors and prevalent CVD [27]. Additional associations between C-reactive protein and ABI, PAD progression, and risk for adverse CVD events among individuals with PAD have been reported [11,13,28]. However, these prior reports were limited as only a few other biomarkers were examined. If we considered each marker separately, both C-reactive protein and fibrinogen were associated with PAD. But in our global model that considered all 12 biomarkers conjointly neither C-reactive protein nor fibrinogen was significantly associated with PAD. One possible explanation may be that the effect of C-reactive protein and fibrinogen may be mediated through interleukin-6 and TNFR2. It is known that interleukin-6 up-regulates both C-reactive protein and fibrinogen and that all three biomarkers are correlated.

4.5. Strengths and limitations

The strengths of the present study include the community-based sample, the simultaneous measurement of a panel of biomarkers, and the direct measurement of clinical factors previously reported to be correlated with PAD and or the inflammatory markers. Several limitations merit comment. The study is cross-sectional and thus we cannot infer that the associations between PAD and the inflammatory markers are causal. We suspect, but cannot establish with the current study design that the relations are bidirectional, with inflammation contributing to PAD and PAD exacerbating systemic inflammation. Conversely, we note that we may have failed to detect small to modest associations. Medication usage (aspirin and lipid lowering treatments) may have altered some inflammatory marker concentrations. Since medication usage was higher in those with PAD, our results may have been biased toward a null result. In addition, the estimated effect sizes of the observed associations were modest; we acknowledge that statistical significance is not synonymous with clinical significance. We acknowledge that walk test data would have enhanced

the accuracy of a PAD diagnosis. Lastly, our sample is primarily white, limiting the ability to generalize our results to other racial and ethnic groups.

5. Conclusions

In a community-based sample interleukin-6 and TNFR2 were significantly associated with PAD accounting for established risk factors. Their effects appear to be independent of each other suggesting that each marker represents a distinct biologic pathway mediating the complex process of vascular inflammation in peripheral atherosclerosis. Further research is needed to establish the role of these markers in predicting incident clinical PAD events and disease progression and to determine whether therapies targeting these markers alter prognosis in patients with PAD.

Conflict of interest

None.

Acknowledgement

The authors appreciate the statistical insights provided by Joseph M. Massaro, PhD.

References

[1] Selvin E, Erlinger TP. Prevalence of and risk factors for peripheral arterial disease in the United States: results from the National Health and Nutrition Examination Survey, 1999–2000. Circulation 2004;110(6):738–43.
[2] O'Hare AM, Katz R, Shlipak MG, Cushman M, Newman AB. Mortality and cardiovascular risk across the ankle-arm index spectrum: results from the Cardiovascular Health Study. Circulation 2006;113(3):388–93.
[3] Resnick HE, Lindsay RS, McDermott MM, Devereux RB, Jones KL, Fabsitz RR, et al. Relationship of high and low ankle brachial index to all-cause and cardiovascular disease mortality: the Strong Heart Study. Circulation 2004;109(6):733–9.
[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(16):1939–42.
[5] Cesari M, Penninx BW, Newman AB, Kritchevsky SB, Nicklas BJ, Sutton-Tyrrell K, et al. Inflammatory markers and onset of cardiovascular events: results from the Health ABC study. Circulation 2003;108(19):2317–22.
[6] Tuomisto K, Jousilahti P, Sundvall J, Pajunen P, Salomaa V. C-reactive protein, interleukin-6 and tumor necrosis factor alpha as predictors of incident coronary and cardiovascular events and total mortality. A population-based, prospective study. Thromb Haemost 2006;95(3):511–8.
[7] Ridker PM, Stampfer MJ, Rifai N. Novel risk factors for systemic atherosclerosis: a comparison of C-reactive protein, fibrinogen, homocysteine, lipoprotein(a), and standard cholesterol screening as predictors of peripheral arterial disease. JAMA 2001;285(19):2481–5.
[8] Ridker PM, Cushman M, Stampfer MJ, Tracy RP, Hennekens CH. Plasma concentration of C-reactive protein and risk of developing peripheral vascular disease. Circulation 1998;97(5):425–8.
[9] Khera A, de Lemos JA, Peshock RM, Lo HS, Stanek HG, Murphy SA, et al. Relationship between C-reactive protein and subclinical atherosclerosis: the Dallas Heart Study. Circulation 2006;113(1):38–43.
[10] Hulthe J, Wikstrand J, Fagerberg B. Relationship between C-reactive protein and intima-media thickness in the carotid and femoral arteries and to antibodies against oxidized low-density lipoprotein in healthy men: the Atherosclerosis and Insulin Resistance (AIR) study. Clin Sci (Lond) 2001;100(4):371–8.
[11] Folsom AR, Pankow JS, Tracy RP, Arnett DK, Peacock JM, Hong Y, et al. Association of C-reactive protein with markers of prevalent atherosclerotic disease. Am J Cardiol 2001;88(2):112–7.
[12] Tracy RP, Psaty BM, Macy E, Bovill EG, Cushman M, Cornell ES, et al. Lifetime smoking exposure affects the association of C-reactive protein with cardiovascular disease risk factors and subclinical disease in healthy elderly subjects. Arterioscler Thromb Vasc Biol 1997;17(10):2167–76.
[13] McDermott MM, Green D, Greenland P, Liu K, Criqui MH, Chan C, et al. Relation of levels of hemostatic factors and inflammatory markers to the ankle brachial index. Am J Cardiol 2003;92(2):194–9.
[14] van dM I, De Maat MP, Hak AE, Kiliaan AJ, del Sol AI, van der Kuip DA, et al. C-reactive protein predicts progression of atherosclerosis measured at various sites in the arterial tree: the Rotterdam Study. Stroke 2002;33(12):2750–5.
[15] Beckman JA, Preis O, Ridker PM, Gerhard-Herman M. Comparison of usefulness of inflammatory markers in patients with versus without peripheral arterial disease in predicting adverse cardiovascular outcomes (myocardial infarction, stroke, and death). Am J Cardiol 2005;96(10):1374–8.
[16] Tzoulaki I, Murray GD, Lee AJ, Rumley A, Lowe GD, Fowkes FG. C-reactive protein, interleukin-6, and soluble adhesion molecules as predictors of progressive peripheral atherosclerosis in the general population: Edinburgh Artery Study. Circulation 2005;112(7):976–83.
[17] Tzoulaki I, Murray GD, Price JF, Smith FB, Lee AJ, Rumley A, et al. Hemostatic factors, inflammatory markers, and progressive peripheral atherosclerosis: the Edinburgh Artery Study. Am J Epidemiol 2006;163(4):334–41.
[18] Hoogeveen RC, Morrison A, Boerwinkle E, Miles JS, Rhodes CE, Sharrett AR, et al. Plasma MCP-1 level and risk for peripheral arterial disease and incident coronary heart disease: Atherosclerosis Risk in Communities study. Atherosclerosis 2005;183(2):301–7.
[19] 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 2002;143(6):961–5.
[20] Jefferson AL, Massaro JM, Wolf PA, Seshadri S, Au R, Vasan RS, et al. Inflammatory biomarkers are associated with total brain volume: the Framingham Heart Study. Neurology 2007;68(13):1032–8.
[21] Abbott RD, McGee DL. The Framingham Study: An Epidemiologic Investigation of Cardiovascular Disease. Section 37. The Probability of Developing Certain Cardiovascular Diseases in Eight Years at Specified Values of Some Characteristics. Bethesda, MD: National Heart, Lung and Blood Institute; 1987.
[22] Benjamin EJ, Dupuis J, Larson MG, Lunetta KL, Booth SL, Govindaraju DR, et al. Genome-wide association with select biomarker traits in the Framingham Heart Study. BMC Med Genet 2007;8(Suppl 1):S11.
[23] Schnabel R, Larson MG, Dupuis J, Lunetta KL, Lipinska I, Meigs JB, et al. Relations of inflammatory biomarkers and common genetic variants with arterial stiffness and wave reflection. Hypertension 2008.
[24] SAS Institute Inc., version 8. Cary, NC: SAS Institute Inc.; 1999.
[25] Signorelli SS, Mazzarino MC, Di Pino L, Malaponte G, Porto C, Pennisi G, et al. High circulating levels of cytokines (IL-6 and TNFalpha), adhesion molecules (VCAM-1 and ICAM-1) and selectins in patients with peripheral arterial disease at rest and after a treadmill test. Vasc Med 2003;8(1):15–9.
[26] Libra M, Signorelli SS, Bevelacqua Y, Navolanic PM, Bevelacqua V, Polesel J, et al. 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 2006;59(2):211–5.
[27] Tzoulaki I, Murray GD, Lee AJ, Rumley A, Lowe GD, Fowkes FG. Inflammatory, haemostatic, and rheological markers for incident peripheral arterial disease: Edinburgh Artery Study. Eur Heart J 2007;28(3):354–62.
[28] McDermott MM, Guralnik JM, Corsi A, Albay M, Macchi C, Bandinelli S, et al. Patterns of inflammation associated with peripheral arterial disease: the InCHIANTI study. Am Heart J 2005;150(2):276–81.
[29] Spranger J, Kroke A, Mohlig M, Hoffmann K, Bergmann MM, Ristow M, et al. Inflammatory cytokines and the risk to develop type 2 diabetes: results of the prospective population-based European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam Study. Diabetes 2003;52(3):812–7.
[30] Sesso HD, Wang L, Buring JE, Ridker PM, Gaziano JM. Comparison of interleukin-6 and C-reactive protein for the risk of developing hypertension in women. Hypertension 2007;49(2):304–10.
[31] Chandrasekharan UM, Mavrakis L, Bonfield TL, Smith JD, DiCorleto PE. Decreased atherosclerosis in mice deficient in tumor necrosis factor-alpha receptor-II (p75). Arterioscler Thromb Vasc Biol 2007;27(3):e16–7.
[32] Chandrasekharan UM, Siemionow M, Unsal M, Yang L, Poptic E, Bohn J, et al. Tumor necrosis factor alpha (TNF-alpha) receptor-II is required for TNF-alpha-induced leukocyte–endothelial interaction in vivo. Blood 2007;109(5):1938–44.
[33] Rodriguez C, Alcudia JF, Martinez-Gonzalez J, Raposo B, Navarro MA, Badimon L. Lysyl oxidase (LOX) down-regulation by TNFalpha: a new mechanism underlying TNFalpha-induced endothelial dysfunction. Atherosclerosis 2007.
[34] Fiotti N, Giansante C, Ponte E, Delbello C, Calabrese S, Zacchi T, et al. Atherosclerosis and inflammation. Patterns of cytokine regulation in patients with peripheral arterial disease. Atherosclerosis 1999;145(1):51–60.

Temporal Trends in Self-Reported Functional Limitations and Physical Disability Among the Community-Dwelling Elderly Population: The Framingham Heart Study

Joanne M. Murabito, MD, ScM, Michael J. Pencina, PhD, Lei Zhu, MA, Margaret Kelly-Hayes, EdD, RN, Peter Shrader, MA, and Ralph B. D'Agostino Sr, PhD

Objectives. We sought to determine change in the prevalence of functional limitations and physical disability among the community-dwelling elderly population across 3 decades.

Methods. We studied original participants of the Framingham Heart Study, aged 79 to 88 years, at examination 15 (1977–1979; 177 women, 103 men), examination 20 (1988–1990; 159 women, 98 men) and examination 25 (1997–1999; 174 women, 119 men). Self-reported functional limitation was defined using the Nagi scale, and physical disability was defined using the Rosow-Breslau and Katz scales.

Results. Functional limitations declined across examinations from 74.6% to 60.5% to 37.9% ($P < .001$) among women and from 54.2% to 37.8% to 27.8% ($P < .001$) among men. Physical disability declined from 74.5% to 48.5% to 34.6% ($P < .001$) among women and 42.3% to 33.3% to 22.8% ($P = .009$) among men. Among women, improvements in functional limitations ($P = .05$) were greater from examination 20 to 25, whereas for physical disability ($P = .02$), improvements were greater from examination 15 to 20. Improvements in function were constant across the 3 examinations in men.

Conclusions. Among community-dwelling elders, the prevalence of functional limitations and physical disability declined significantly in both women and men from the 1970s to the 1990s. This may in part be due to improvements in technological devices used to maintain independence. Further work is needed to identify the underlining causes of the decline so preventative measures can be established that promote independence for the elderly population. (*Am J Public Health.* 2008;98:1256–1262. doi:10.2105/AJPH.2007.128132)

National surveys and epidemiological studies have reported a significant decline in self-reported functional limitations and physical disability among older adults.[1–7] Despite consensus among reports, uncertainty exists with regard to the magnitude, rate, and specific characteristics of the disability decline.[5,6,8] Variations in study samples, evolving measures of functional limitation and disability, and differences in study questions and responses contribute to the inconsistencies in disability trends.[5,6] Furthermore, disparities exist in the improvement in function, with marked variations according to age, gender, race, and socioeconomic and educational attainment.[1,9,10] Compared with men, women report greater difficulty with physical function and less recovery from disability.[1] Surveys have reported that declines in functional limitations occurred only among women[11] or were larger among women than among men,[12] whereas others note that disability declines were about the same among women and men.[3,13] Thus, it remains unclear if the disability gaps between men and women have narrowed or remained stable over time.[4]

The causes for the improved disability trends are not well understood. One possible explanation is the "compression of morbidity" hypothesis, whereby disease and disability are postponed until the end of the lifespan.[14,15] However, the consequences of an increase in life expectancy in the United States in relation to the overall health of older adults continue to be debated. Other divergent paradigms have been proposed to describe the possible health-related consequences of living longer, including a rise in chronic disease and disability[16] and a dynamic equilibrium whereby declines in mortality result in increases in chronic disease with lesser severity and disability.[17,18]

We obtained self-reported information on functional limitations and physical disability among surviving members of the original cohort of the Framingham Heart Study in late life (aged 79–88 years) who attended research examinations over 3 points in time from the 1970s to the 1990s. We hypothesized that the prevalence of functional limitations and physical disability would decline over time among elders, with a greater decline among women than among men. Our study cohort is particularly well suited for this investigation, because the Framingham Disability Study[19] introduced questionnaires to measure self-reported functional limitations and physical disability beginning in 1976 that were repeated on successive examinations. Moreover, this cohort has been well characterized for over 50 years, with documentation of validated medical conditions and measurement of risk factors.

METHODS

Study Setting and Participants

The Framingham Heart Study was initiated in 1948 when 5209 participants, aged 28 to 62 years, were enrolled in a prospective cardiovascular disease study.[20,21] Since the study inception, participants have been examined biennially, including a standardized physician-administered medical history and physical examination, electrocardiogram, noninvasive testing, and measurement of lipids and glucose. Written informed consent was obtained from each of the participants.

Because we were interested in studying trends in late-life disability, we restricted our

study sample to participants aged 79 to 88 years. The age restriction ensured age comparability across time points, minimizing any confounding effects of age and permitting adequate numbers of both men and women at a given age across the time points. We did not study younger ages, because improvements in disability have been reported for adults aged 55 to 70 years.[22] Participants with dementia were excluded to enhance the accuracy of the disability data, because the data are self-reported. Original cohort examinations 15 (August 1977–November 1979), 20 (January 1988–June 1990), and 25 (October 1997–November 1999) were chosen for study, because examinations 14 and 15 were the first time functional data were collected and the subsequent examinations were conducted at approximately 10-year intervals. Finally, collection of physical function data at examination 15 was limited to noninstitutionalized participants; therefore, we restricted our sample to participants attending a clinic examination. Hence, our final study sample consisted of 3 different groups of participants, aged 79 to 88 years, 1 group for each chosen examination time point. (An additional table of data on the original Framingham cohort attending index examinations 15, 20, and 25 provides the details of eligibility and exclusion for the final study sample and is available as a supplement to the online version of this article at http://www.ajph.org.)

Assessment of Functional Limitations and Physical Disability

In our study, we defined functional limitations based on the physical performance scale adapted from Nagi[20] and physical disability with the modified Katz Activity of Daily Living (ADL)[22] scale and the Rosow-Breslau functional health scale.[21] These self-reported measures have been used in other large population-based studies with high test–retest reliability, permitting use in longitudinal analyses.[1,23–27] Technicians interviewed each participant by asking standardized questions and recording response choices at each examination. The script used for the Katz ADL scale was modified from examination 15 ("Other than when you might have been in the hospital, was there any time during the past 12 months in which you needed help from another person or from some special equipment or device?") to exams 20 and 25 ("During the course of a normal day, can you do the following activities independently or do you need human assistance or use of a device?"). Functional limitation was considered present if a participant reported more than a little difficulty on any of the 7 items of the Nagi scale: pulling or pushing large objects, like a living room chair; either stooping, crouching, or kneeling; reaching or extending arms above shoulder level; reaching or extending arms below shoulder level; either writing or handling or fingering small objects; standing in one place for long periods; and sitting for at least 1 hour. Response choices included no difficulty, a little difficulty, some difficulty, a lot of difficulty, or don't do under doctor's orders and unable to do (examinations 20 and 25).

For physical disability determination, we used the Rosow-Breslau functional health scale to document the following 3 gross-mobility tasks: walk 0.5 mile, walk up and down stairs to second floor, and do heavy work around the house. Participants reported whether they were able or unable to do these tasks without help. At examination 25, the question of walking up and down 1 flight of stairs was asked with the modified Katz ADL scale. The modified Katz ADL scale included the following 5 items: bathing, dressing, eating, getting from bed to chair, and walking across a small room. On examinations 20 and 25, getting from bed to chair was changed to transferring (getting in and out of a chair), and walking across a small room was changed to walking on a level surface for about 50 yards. Participants reported the following: no help needed (independent), needed help from special equipment or device, needed help from another person, or needed help from both a person and special equipment. Disability was defined at each examination as requiring human assistance. If the participant reported not performing the activity during a normal day, then the response was set to missing.

Chronic Medical Conditions and Health Behaviors

Hypertension was defined as a blood pressure of 140/90 mm Hg or greater or use of antihypertensive medication. Body mass index was calculated as weight in kilograms divided by height in meters squared. Diabetes was defined by a casual blood glucose of 200 mg/dL or greater, or use of insulin or oral hypoglycemic agents. An end point committee comprised of 3 senior investigators (or a panel of study neurologists) adjudicated cardiovascular outcomes with all available medical records, employing standardized criteria in place since study inception.[28] The Center for Epidemiological Studies depression scale (CES-D) was administered to participants at examinations 22 and 25. Depressive symptoms were considered present if the CES-D score was 16 or higher.[29]

A current cigarette smoker was defined as regular smoking in the year preceding the examination. Participants were asked if they drank beer, wine, or liquor at least once per month, and the number of drinks on an average week was recorded. Chronic obstructive pulmonary disease (COPD) was considered present if the ratio of the forced expiratory volume at 1 second per forced vital capacity was less than 70% of predicted. At examination 25, spirometry was not performed and the diagnosis of COPD was defined by the physician opinion of chronic bronchitis or chronic symptoms (cough, sputum production). Marital status was updated at all examinations.

Statistical Analysis

The gender-specific prevalence of chronic medical conditions, health behaviors, sociodemographic characteristics, as well as the prevalence of functional limitation or physical disability for each item of the Nagi, Rosow-Breslau, and modified Katz scales were calculated as mean value for continuous variables and percentage for dichotomous variables at each examination studied. Next, we conducted gender-specific analyses adjusted for age, with examination (15, 20, and 25) as the exposure variable, to investigate time trends in functional limitations and physical disability. Analyses were as follows: (1) linear regression (PROC GLM in SAS[30]) was used to calculate the least square means and the 95% confidence intervals for the number of items on the Nagi scale reported as a limitation, and on the number of items on the Rosow-Breslau and modified Katz scale reported as an impairment; and (2) logistic regression (PROC

LOGISTIC in SAS) was used to calculate the proportion of participants reporting a functional limitation and physical disability. We also created a summary measure of functional limitations and physical disability, taking into account all items of the 3 scales.

Next, we set out to determine whether the magnitude of absolute change in functional limitations and physical disability varied between men and women and between examination periods (examination 15 through examination 20 vs examination 20 through examination 25) by comparing differences in mean numbers of items with limitation or impairment (PROC GLM, *z* tests). Because women were noted to report higher levels of functional limitation and physical disability than were men, the absolute decline and, thus, improvement in function and disability might be expected to be greater among women. We therefore tested for any difference in the relative decline in functional limitation and physical disability between men and women by using asymptotic normal theory applied to gender-specific logistic regression slopes. All analyses were conducted with SAS/STAT version 9.1.[30]

RESULTS

Chronic medical conditions, health behaviors, and sociodemographic characteristics of the sample at each examination are shown in Table 1. The prevalence of obesity increased across examinations, particularly among men, as did the prevalence of cardiovascular disease (men only) and cancer, whereas the prevalence of COPD declined in concert with the decline in prevalence of current cigarette smoking. The prevalence of at least 1 health condition remained constant at about

TABLE 1—Characteristics of Community-Dwelling Elderly Men and Women: Original Cohort, Framingham Heart Study, Framingham, Massachusetts, 1977–1999

	Women			Men		
	Examination 15 (1977–1979)	Examination 20 (1988–1990)	Examination 25 (1997–1999)	Examination 15 (1977–1979)	Examination 20 (1988–1990)	Examination 25 (1997–1999)
Total, no.	177	159	174	103	98	119
Chronic medical conditions						
Mean age, y	82.2	81.9	82.2	81.6	82.1	82.3
Hypertension,[a] %	76.1	81.1	83.3	56.4	85.7	76.5
Hypertension treatment, %	47.9	63.5	60.3	33.0	61.2	55.9
Total cholesterol ≥ 240 mg/dL or Rx, %	38.3	24.5	32.8	12.9	11.7	22.9
Diabetes,[b] %	9.8	8.2	5.7	7.4	13.3	10.1
Obesity (BMI ≥ 30 kg/m^2), %	12.4	15.5	20.7	2.2	12.4	15.3
Cardiovascular disease,[c] %	37.9	28.3	29.3	41.7	44.9	55.5
Cancer, %	11.3	14.5	19.0	15.5	16.3	24.4
Chronic obstructive pulmonary disease, %	9.9	9.6	5.7	18.5	12.5	10.3
Hip fracture, %	5.1	8.8	6.9	1.9	1.0	1.7
Depressive symptoms,[d] %	NA	16.7	10.4	NA	6.1	7.0
At least 1 medical condition,[e] %	52.7	52.2	51.1	63.0	58.7	75.2
Health behaviors						
Current cigarette smoker, %	4.9	5.7	4.0	15.3	4.1	5.1
Alcohol intake, no. drinks/wk	1.9	3.6	2.5	6.1	6.4	5.4
Marital status, %						
Married	17.8	19.0	30.2	73.4	80.9	73.7
Widowed	69.9	63.3	59.3	22.3	14.9	21.2
Single	10.4	12.0	8.7	2.1	2.1	3.4
Divorced	1.9	5.7	1.8	2.2	2.1	1.7
Education, high school or greater, %	53.2	58.3	78.5	46.5	54.3	76.3
Living situation, lives alone, %	NA	72.2	46.2	NA	17.3	19.3
Work status, currently working, %	2.3	6.3	6.3	12.6	14.3	16.0
Subjective health, good or excellent, %	66.9	80.5	79.2	80.0	76.5	84.0

Note. BMI = body mass index; NA = not available, data not collected; Rx = medication treatment.
[a]Hypertension was defined as blood pressure at 140/90 mm Hg or greater or on antihypertensive medication treatment.
[b]Diabetes was defined as causal blood glucose of 200 mg/dL or greater or treatment with oral hypoglycemic agent or insulin.
[c]Cardiovascular disease included coronary heart disease, stroke, transient ischemic attack, congestive heart failure, or intermittent claudication.
[d]Depressive symptoms were assessed at examinations 22 and 25 with the Center for Epidemiological Studies Depression scale. A score of 16 or greater defined the presence of depressive symptoms.
[e]Medical conditions included any of the following 5 conditions: diabetes, cardiovascular disease, cancer, chronic obstructive pulmonary disease, and hip fracture.

TABLE 2—Self-Reported Prevalence of Functional Limitation or Physical Disability Among Community-Dwelling Elderly Men and Women: Original Cohort, Framingham Heart Study, Framingham, Massachusetts, 1977–1999

	Women			Men		
Scale and Scale Items	Examination 15 (n = 177), %	Examination 20 (n = 159), %	Examination 25 (n = 174), %	Examination 15 (n = 103), %	Examination 20 (n = 98), %	Examination 25 (n = 119), %
Nagi physical performance scale						
Pulling or pushing large objects	40.6	31.6	18.3	14.7	11.2	6.7
Either stooping, crouching, or kneeling	48.3	47.8	18.5	23.5	18.4	17.6
Reaching or extending arms above shoulder level	9.7	14.6	7.5	5.9	8.2	4.2
Reaching or extending arms below shoulder level	4.5	3.8	1.1	1.0	2.0	0.8
Writing, handling, or fingering small objects	17.5	11.9	5.7	15.5	11.2	10.1
Standing in place for long periods say 15 min	39.0	29.7	15.9	22.3	16.3	12.2
Sitting for at least 1 hour	9.7	5.0	4.0	9.9	6.1	0.8
Rosow-Breslau functional health scale						
Heavy work around the house	65.5	39.0	30.5	35.9	31.6	19.3
Walks up and down stairs to second floor[a]	26.0	7.5	6.9	6.8	2.0	0.0
Walk one half mile	39.5	25.2	16.7	16.5	11.2	14.3
Modified Katz Activities of Daily Living scale						
Bathing/personal grooming	15.3	2.5	2.9	4.9	2.0	0.0
Dressing	1.7	1.9	3.4	2.9	3.1	0.8
Eating	0.0	0.0	1.1	0.0	2.0	0.8
Getting from a bed to a chair[b]	0.0	1.3	2.9	1.0	3.1	0.0
Walking across small room[b]	4.5	1.3	0.6	0.0	1.0	0.8
Any special equipment or device use	7.9	13.2	28.2	3.9	4.1	17.6

Note. The mean age of the cohort was 82 years. Disability was defined as human assistance needed, minimally dependent, or dependent on the modified Katz Activities of Daily Living Scale and as unable to do so on the Rosow-Breslau functional health scale. Using the Nagi physical performance scale a functional limitation was defined as some difficulty, a lot of difficulty, unable to do, or do not do on physician order.

[a]At examination 25, this question was asked as part of the modified Katz Activities of Daily Living scale.

[b]At examinations 20 and 25, the question was changed to "transferring (getting in and out of a chair)" and "walking on a level surface about 50 yards."

50% for women, whereas among men, the prevalence of at least 1 health condition changed across the 3 examination periods, from 63.0% to 58.7% to 75.2%. Striking differences in marital status among men and women were noted. Thus, less than 20% of men reported living alone, and 72.2% of women at examination 20 and 46.2% of women at examination 25 lived alone. Most participants reported their health to be good or excellent.

Women reported greater functional limitations and physical disability for almost all items of the Nagi physical performance scale, the Rosow-Breslau functional health scale, and the modified Katz ADL scale at all examinations (Table 2). Of note, the magnitude of the difference between men and women in the self-reported prevalence of disability or functional limitation for some scale items narrowed over time. For example, at examination 15, 39.5% of women and 16.5% of men reported inability to walk 0.5 miles; however, at examination 25, 16.7% of women and 14.3% of men were unable to do so. The use of special equipment or devices increased across exams among both women and men.

Gender-specific trends in functional limitations and physical disability adjusted for age are shown in Table 3. A significant decline in functional limitations and disability was observed in both women and men. For example, at examination 15, 74.6% of women reported at least 1 functional limitation on the Nagi physical performance scale compared with 60.5% of women at examination 20 and 37.9% of women at examination 25 ($P<.001$). Corresponding reports among men at examinations 15, 20, and 25 were 54.2%, 37.8%, and 27.8% ($P<.001$), respectively. Likewise, self-reported physical disability assessed with the Rosow-Breslau and modified Katz ADL scales declined across the 3 examination time periods, from 74.5% to 48.5% to 34.6% ($P<.001$), respectively, among women and 42.3% to 33.3% to 22.8% ($P=.009$) among men. The mean number of scale items reported with a limitation or impairment also decreased across exams among both men and women. By examination 25, 63.6% of men and 50.9% of women reported that they were free of any functional limitations and physical disability. Repeating the analyses, adjusting for chronic medical conditions defined by the presence or absence of cardiovascular disease, cancer, diabetes, and hip fracture with a score from 0 to 4 did not change the trends.

Next, we examined whether the magnitude of the decline in limitations and physical disability differed between men and women and whether the decline differed across time, comparing the change that occurred between examinations 15 and 20 to that that occurred between examinations 20 and 25 in terms of the number of reported scale items with limitation or impairment. Women started with

TABLE 3—Gender-Specific Trends in Self-Reported Functional Limitations and Physical Disability Among Community-Dwelling Elderly Men and Women, Adjusted for Age: Original Cohort, Framingham Heart Study Framingham, Massachusetts, 1977–1999

	Women				Men			
Physical Function Scale	Examination 15 (n = 177)	Examination 20 (n = 159)	Examination 25 (n = 174)	*P*	Examination 15 (n = 103)	Examination 20 (n = 98)	Examination 25 (n = 119)	*P*
Functional limitation, Nagi scale								
No. of items with limitation, mean (95% CI)	1.67 (1.46, 1.88)	1.45 (1.23, 1.67)	0.70 (0.48, 0.91)	<.001	0.93 (0.71, 1.14)	0.73 (0.51, 0.96)	0.53 (0.32, 0.73)	.03
Any difficulty on the scale, % (95% CI)	74.6 (66.7, 81.1)	60.5 (51.5, 68.8)	37.9 (29.9, 46.5)	<.001	54.2 (43.1, 65.0)	37.8 (27.6, 49.2)	27.8 (19.6, 37.9)	<.001
Physical disability, Rosow–Breslau, and Katz ADL scales, mean (95% CI)								
No. of items with impairment, mean (95% CI)	1.52 (1.34, 1.70)	0.81 (0.62, 1.00)	0.64 (0.46, 0.82)	<.001	0.71 (0.52, 0.90)	0.55 (0.36, 0.75)	0.35 (0.17, 0.52)	.02
Any difficulty on the scales, % (95% CI)	74.5 (66.7, 80.9)	48.5 (39.8, 57.2)	34.6 (27.0, 43.0)	<.001	42.3 (31.9, 53.3)	33.3 (23.7, 44.4)	22.8 (15.4, 32.3)	.009
Summary, all 3 scales								
No. of items with limitation or impairment, mean (95% CI)	3.19 (2.86, 3.52)	2.26 (1.91, 2.61)	1.34 (1.00, 1.67)	<.001	1.64 (1.28, 2.00)	1.29 (0.92, 1.66)	0.87 (0.54, 1.20)	.009
Any difficulty on the scales, % (95% CI)	88.1 (81.9, 92.4)	69.6 (60.8, 77.0)	49.1 (40.4, 57.9)	<.001	63.9 (52.9, 73.5)	48.8 (37.7, 59.8)	36.4 (27.1, 46.9)	<.001

Note. ADL = Activity of Daily Living; CI = confidence interval. Analysis adjusted for age.

greater limitations and disability compared with men (Table 3) and experienced a greater absolute decline in both functional limitations (*P*=.008) and physical disability (*P*=.005) than did men. However, the relative difference in the decline determined by examining the proportion of women versus men reporting any difficulty on the scale was significantly different only for physical disability (*P*=.03). Absolute improvements in both functional limitations and physical disability were constant across the examination time periods (examinations 15 through 20 and examinations 20 through 25) among men. Among women, there was a greater improvement in functional limitation between examinations 20 and 25 compared with that between examinations 15 and 20 (*P*=.05), whereas the improvement in physical disability was more marked between examinations 15 and 20 compared with that between examinations 20 and 25 (*P*=.02).

DISCUSSION

In our sample of community-dwelling elderly people, we found a significant decline in self-reported functional limitations and physical disability in both women and men over 3 examinations occurring from the late 1970s to the late 1990s. At examination 25 (1997–1999), more than half of men and women were free of both functional limitations and disability. This finding suggests continued progress when compared with a report by Liao et al., in which 42% of men and 34% of women 70 years and older were found to be physically robust, without any limitations or disabilities.[31]

In our study, women reported a greater burden of functional limitations and physical disability than did men. Hence, the absolute decline in both limitations and disability was significantly greater among women than among men; however, the relative difference in the magnitude of decline was significantly different between men and women only for physical disability. Moreover, the absolute improvement in function and disability was constant over the examinations in men, whereas among women, the improvement in disability was greater between examinations 15 and 20 (late 1970s to late 1980s) than between examinations 20 and 25 (late 1980s to late 1990s). These findings are in contrast to national survey data, which suggest that the disability decline has accelerated in more recent years.[32,33] In accordance with those reports, the magnitude of improvement in functional limitations among women was greatest at more recent examinations. The decline in disability in our study may, in part, be related to the notable increase in the use of special equipment and devices that facilitate greater independence. This finding is consistent with other reports that noted an increase in the proportion of the community-dwelling elderly population who used equipment to bathe.[5] Our work highlights the importance of studying gender-specific trends in disability, as well as the need for careful attention to the specific measures used to define disability to determine whether any improvements include all types of limitations and impairments.

It is noteworthy that, among men, the decline in physical disability and improvement in functional limitations occurred despite an increase in the prevalence of chronic medical conditions, whereas among women, the improvements occurred in the absence of a change in the prevalence of chronic medical conditions. Our findings in men may be related to improvements in diagnosis and treatment of chronic conditions that occurred over time. Successful prevention and postponement of disablement in the elderly depend, in part, on efforts at early diagnosis of illness and subsequent focused interventions.[34] As in our study, national survey data demonstrated an increase in self-reported medical conditions over 2 points in time (1984 and 1994), yet many of the conditions had less of a debilitating effect.[18] We extended this knowledge by

focusing on older adults (mean age=82 years), including not only functional limitations, but also physical disability measures. In addition, rather than relying on self-reported medical conditions, the conditions in our study were directly measured or validated with medical records. We acknowledge that the number of medical conditions in our study was not exhaustive. Ferrucci et al. have raised the hypothesis that there may be gender-related differences in the lifetime prevalence of lethal versus disabling diseases.[35] Hence, it is possible that, among women, important disabling conditions not included in our study have become less debilitating over time.[18]

The 2001 World Health Organization International Classification of Functioning, Disability, and Health highlights the importance of environmental and personal factors in the disablement process. For older adults, disability generally refers to the ability to live independently and perform self-care activities. In our sample, the prevalence of self-reported mobility disability was significant among both women and men even at the most recent examination. Gross mobility is often the first area in which older adults report difficulty,[24,36] yet little research has been done to determine how environmental factors influence the process and trajectory of disability.[37] The rise in reported special equipment or device use likely contributed to the improvement; however, other environmental changes, such as home modifications, may have positively influenced the trends. Personal factors, such as gender, age, education, lifestyle habits, and marital status, may also play a role in disability.[38] It is notable that most women in our sample were widowed and reported living alone. Prior work has shown the importance of family and social factors to risk of institutionalization after stroke.[39] These same factors may be operational in the disablement process.

Limitations

Our study has several limitations that merit comment. We focused on community-dwelling, nondemented elders, and included only participants attending an on-site clinic examination, because participants were not offered examinations in their personal residence or nursing home at the start of this study (examination 15). The proportion of institutionalized elderly declined during the years of our study.[32] This trend would have resulted in an increase in persons with disability in the community and biased our results toward the null.

All individuals in our sample were White, and thus our results may not pertain to other racial or ethnic groups. The decline in disability in recent years was reported to be greater among the Black population than among the non-Black population.[32] Additionally, our sample is fairly well educated; three quarters of participants at examination 25 had attained a high school or greater education. Educational achievement has been consistently linked to longevity[40] and improvements in late-life function.[4,8,13]

Performance-based measures were not included in the examinations studied for this report. Conceptualizations of disability and individual perceptions of social roles, especially for women, have evolved over time. It is unclear how much of the decline in disability among women in our sample is caused by changing self-perceptions of ability to perform tasks (social desirability) versus other factors, such as innovations in diagnosis and treatment of chronic illness, improvements in health-related behaviors, especially smoking cessation and increased physical activity, and the emergence of alternative living arrangements and expanded use of assistive devices allowing older persons to maintain independence.[6] Finally, we acknowledge that change in the wording of the scale items and response choices may have contributed to the changes in self-reported functional limitations and physical disability in our study.

Conclusions

We found that for community-dwelling men and women aged 79 to 88 years, the prevalence of functional limitations and physical disability declined significantly over 3 examinations from the late 1970s to the late 1990s. The relative magnitude of the decline in physical disability was greater among women than among men. In contrast to the acceleration in the disability decline in recent years noted in national survey data, the decline in limitations and disability was constant over time in men in our sample, whereas among women, the decline was greatest in the earlier part of our study (examination 15 to examination 20, late 1970s to late 1980s). The improvement in physical function trends was noted in concert with a marked increase in reported use of special equipment and devices used to maintain independence. Future work is needed to determine the underlying causes contributing to the declines in limitations and disability in old age so that preventative measures can be put in place to promote and maintain independence until the end of life. It remains uncertain whether the improvements in functional limitations and physical disability will continue, given the unfavorable direction of the prevalence of obesity and physical activity in the general population, factors known to predict incident disability[41] and declines in physical performance.[42] ■

About the Authors

Joanne M. Murabito, Michael J. Pencina, Margaret Kelly-Hayes, and Ralph B. D'Agostino Sr are with the National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, MA. Joanne M. Murabito is with the Department of Medicine, Section of General Internal Medicine, Boston University School of Medicine, Boston, MA. Margaret Kelly-Hayes is with the Department of Neurology, Boston University School of Medicine, Boston. Michael J. Pencina, Lei Zhu, Peter Shrader, and Ralph B. D'Agostino Sr are also with the Department of Mathematics and Statistics, Boston University, Boston.

Requests for reprints should be sent to Joanne M. Murabito, MD, ScM, 73 Mount Wayte Ave, Suite 2, Framingham, MA 01702-5827 (e-mail: murabito@bu.edu).

This article was accepted February 10, 2008.

Contributors

J.M. Murabito conceptualized and designed the study, acquired, analyzed, and interpreted the data, and drafted and revised the article for important intellectual content. M.J. Pencina conceptualized and designed the study, analyzed and interpreted the data, and revised the article for important intellectual content. L. Zhu analyzed and interpreted the data. M. Kelly-Hayes conceptualized and designed the study, acquired, analyzed, and interpreted the data, and revised the article for important intellectual content. P. Shrader analyzed and interpreted the data. R.B. D'Agostino Sr conceptualized and designed the study and acquired, analyzed, and interpreted the data.

Acknowledgments

The Framingham Heart Study is supported by a National Institues of Health (NIH) National Heart, Lung, and Blood Institute contract (N01-HC-25195). M. Kelly-Hayes was additionally supported by the NIH/National Institute on Aging (grant 5R01-AG08122), Epidemiology of Dementia in the Framingham Study, and NIH/National Institute of Neurological Disorders and Stroke (grant 5R01-NS17950; Precursors of Stroke Incidence and Prognosis).

Human Participant Protection

The institutional review board of Boston University Medical Center approved the content of each Framingham Heart Study examination.

References

1. Beckett LA, Brock DB, Lemke JH, et al. Analysis of change in self-reported physical function among older persons in four population studies. *Am J Epidemiol.* 1996;143:766–778.

2. Costa DL. Changing chronic disease rates and long-term declines in functional limitation among older men. *Demography.* 2002;39:119–137.

3. Crimmins EM, Saito Y, Reynolds SL. Further evidence on recent trends in the prevalence and incidence of disability among older Americans from two sources: the LSOA and the NHIS. *J Gerontol B Psychol Sci Soc Sci.* 1997;52:S59–S71.

4. Freedman VA, Martin LG, Schoeni RF. Recent trends in disability and functioning among older adults in the United States: a systematic review. *JAMA.* 2002; 288:3137–3146.

5. Freedman VA, Crimmins E, Schoeni RF, et al. Resolving inconsistencies in trends in old-age disability: report from a technical working group. *Demography.* 2004;41:417–441.

6. *Disability Decline; What We Know; and What We'd Like To Know.* National Institute on Aging-National Bureau of Economic Research workshop summary. Bethesda, MD: National Institutes of Health Stone House; 2001.

7. *Federal Interagency Forum on Aging-Related Statistics. Older Americans 2004: Key Indicators of Well-Being. Federal Interagency Forum on Aging-Related Statistics.* Washington, DC: US Government Printing Office. November 2004. Available at: http://www.agingstats.gov/Agingstatsdotnet/Main_Site/Data/2004_Documents/entire_report.pdf. Accessed March 7, 2008.

8. Waidmann TA, Liu K. Disability trends among elderly persons and implications for the future. *J Gerontol B Psychol Sci Soc Sci.* 2000;55:S298–S307.

9. Minkler M, Fuller-Thomson E, Guralnik JM. Gradient of disability across the socioeconomic spectrum in the United States. *N Engl J Med.* 2006;355:695–703.

10. Crimmins EM. Trends in the health of the elderly. *Annu Rev Public Health.* 2004;25:79–98.

11. Crimmins E, Saito Y. Change in the prevalence of diseases among older Americans: 1984–1994. *Demogr Res.* [serial online] 2000;3:article 9. Available at: http://www.demographic-research.org/Volumes/Vol3/9/3-9.pdf. Accessed March 7, 2008.

12. Freedman VA, Martin LG. Understanding trends in functional limitations among older Americans. *Am J Public Health.* 1998;88:1457–1462.

13. Schoeni RF, Freedman VA, Wallace RB. Persistent, consistent, widespread, and robust? Another look at recent trends in old-age disability. *J Gerontol B Psychol Sci Soc Sci.* 2001;56:S206–S218.

14. Fries JF. Aging, natural death, and the compression of morbidity. *N Engl J Med.* 1980;303:130–135.

15. Fries JF. Measuring and monitoring success in compressing morbidity. *Ann Intern Med.* 2003;139 (5 pt 2):455–459.

16. Gruenberg EM. The failures of success. *Milbank Mem Fund Q Health Soc.* 1977;55:3–24.

17. Manton KG. Changing concepts of morbidity and mortality in the elderly population. *Milbank Mem Fund Q Health Soc.* 1982;60:183–244.

18. Freedman VA, Martin LG. Contribution of chronic conditions to aggregate changes in old-age functioning. *Am J Public Health.* 2000;90:1755–1760.

19. Jette AM, Branch LG. The Framingham Disability Study: II. Physical disability among the aging. *Am J Public Health.* 1981;71:1211–1216.

20. Dawber TR, Meadors GF, Moore FE Jr. Epidemiological approaches to heart disease: the Framingham Study. *Am J Public Health.* 1951;41:279–281.

21. Dawber TR, Kannel WB, Lyell LP. An approach to longitudinal studies in a community: the Framingham Study. *Ann N Y Acad Sci.* 1963;107:539–556.

22. Allaire SH, LaValley MP, Evans SR, et al. Evidence for decline in disability and improved health among persons aged 55 to 70 years: the Framingham Heart Study. *Am J Public Health.* 1999;89:1678–1683.

23. Smith LA, Branch LG, Scherr PA, et al. Short-term variability of measures of physical function in older people. *J Am Geriatr Soc.* 1990;38:993–998.

24. Alexander NB, Guire KE, Thelen DG, et al. Self-reported walking ability predicts functional mobility performance in frail older adults. *J Am Geriatr Soc.* 2000;48:1408–1413.

25. Crawford SL, Jette AM, Tennstedt SL. Test-retest reliability of self-reported disability measures in older adults. *J Am Geriatr Soc.* 1997;45:338–341.

26. Jette AM. The Functional Status Index: reliability and validity of a self-report functional disability measure. *J Rheumatol Suppl.* 1987; 14 Suppl 15:15–21.

27. Cronin-Stubbs D, de Leon CF, Beckett LA, Field TS, Glynn RJ, Evans DA. Six-year effect of depressive symptoms on the course of physical disability in community-living older adults. *Arch Intern Med.* 2000;160: 3074–3080.

28. Abbott RD. *The Framingham Study: An Epidemiologic Investigation of Cardiovascular Disease, Section 37: The Probability of Developing Certain Cardiovascular Diseases in Eight Years at Specified Values of Some Characteristics.* Bethesda, MD: National Heart, Lung and Blood Institute; 1987.

29. Irwin M, Artin KH, Oxman MN. Screening for depression in the older adult: criterion validity of the 10-item Center for Epidemiological Studies Depression scale (CES-D). *Arch Intern Med.* 1999;159:1701–1704.

30. SAS/STAT software, version 8.2. *SAS System for Windows.* Cary, NC: SAS Institute Inc; 1999.

31. Liao Y, McGee DL, Cao G, Cooper RS. Recent changes in the health status of the older US population: findings from the 1984 and 1994 supplement on aging. *J Am Geriatr Soc.* 2001;49:443–449.

32. Manton KG, Gu X. Changes in the prevalence of chronic disability in the United States Black and non-Black population above age 65 from 1982 to 1999. *Proc Natl Acad Sci U S A.* 2001;98:6354–6359.

33. Fries JF. Reducing disability in older age. *JAMA.* 2002;288:3164–3166.

34. Laukkanen P, Sakari-Rantala R, Kauppinen M, Heikkinen E. Morbidity and disability in 75- and 80-year-old men and women. A five-year follow-up. *Scand J Soc Med Suppl.* 1997;53:79–106.

35. Ferrucci L, Turchi A, Fumagalli S, et al. Sex-related differences in the length of disability prior to death in older persons. *Aging Clin Exp Res.* 2003;15:310–314.

36. Fried LP, Bandeen-Roche K, Chaves PH, Johnson BA. Preclinical mobility disability predicts incident mobility disability in older women. *J Gerontol A Biol Sci Med Sci.* 2000;55:M43–M52.

37. Shumway-Cook A, Patla A, Stewart A, Ferrucci L, Ciol MA, Guralnik JM. Environmental components of mobility disability in community-living older persons. *J Am Geriatr Soc.* 2003;51:393–398.

38. Field MJ, Jette AM, Martin L, eds. *Workshop in Disability in America: A New Look–Summary and Background Papers.* Washington, DC: National Academies Press; 2006. Available at: http://www.nap.edu/catalog/11579.html. Accessed September 26, 2008.

39. Kelly-Hayes M, Wolf PA, Kannel WB, Sytkowski P, D'Agostino RB, Gresham GE. Factors influencing survival and need for institutionalization following stroke: the Framingham Study. *Arch Phys Med Rehabil.* 1988; 69:415–418.

40. 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;53:1944–1950.

41. Boyle PA, Buchman AS, Wilson RS, Bienias JL, Bennett DA. Physical activity is associated with incident disability in community-based older persons. *J Am Geriatr Soc.* 2007;55:195–201.

42. Seeman TE, Charpentier PA, Berkman LF, et al. Predicting changes in physical performance in a high-functioning elderly cohort: MacArthur studies of successful aging. *J Gerontol.* 1994;49:M97–M108.

PERSPECTIVES

Evidence Does Not Support Clinical Screening of Literacy

Michael K. Paasche-Orlow, MD, MA, MPH[1] *and Michael S. Wolf, PhD, MPH*[2]

[1]Section of General Internal Medicine, Boston University School of Medicine, Boston, MA, USA; [2]Health Literacy and Learning Program, Institute for Healthcare Studies and Division of General Internal Medicine, Northwestern University, Chicago, IL, USA.

Limited health literacy is a significant risk factor for adverse health outcomes. Despite controversy, many health care professionals have called for routine clinical screening of patients' literacy skills. Whereas brief literacy screening tools exist that with further evaluation could potentially be used to detect limited literacy in clinical settings, no screening program for limited literacy has been shown to be effective. Yet there is a noted potential for harm, in the form of shame and alienation, which might be induced through clinical screening. There is fair evidence to suggest that possible harm outweighs any current benefits; therefore, clinical screening for literacy should not be recommended at this time.

KEY WORDS: literacy; health literacy; screening; clinical care; communication; patient.
J Gen Intern Med 23(1):100–2
DOI: 10.1007/s11606-007-0447-2

The relationship between limited literacy and adverse health outcomes has been well documented,[1] and seminal reports about the "problem of health literacy" have been issued by the Institute of Medicine,[2] Agency for Healthcare Research and Quality,[3] American Medical Association,[4] and Joint Commission on the Accreditation of Hospital Organizations[5] among others. Whereas these reports all recognize the growing need to establish an effective response within health care systems to address the problem, few intervention strategies have been studied.[6]

Despite the lack of available interventions, many health care professionals have called for clinical screening of literacy, and brief screening tools have already been developed for this purpose.[7–9] Yet only minimal direct evidence is currently available evaluating the potential benefit of literacy screening. A single study by Seligman and colleagues examined the efficacy of clinical screening for improving physicians' care management strategies among diabetic patients.[10] Whereas physicians who received notification of their patients' health literacy level were more likely to use supportive strategies, such as involving family members, they were less satisfied with patient visits than physicians not receiving notifications, and screening ultimately exhibited no benefit for patients.

Received June 14, 2007
Revised October 18, 2007
Accepted October 24, 2007
Published online November 9, 2007

Whereas screening for limited literacy might sound appealing to mitigate the health impact of this prevalent, dangerous, and often silent phenomenon, there are several important considerations that must first be addressed. We present a critical review of the case for literacy screening in clinical settings that summarizes the utility of literacy screening to date and its value added, if any, to medical care.[11]

RATIONALE FOR CLINICAL SCREENING

A call for literacy screening is driven by the preponderance of research demonstrating associations and the predictive power of literacy skills on various health outcomes. Specifically, studies have shown that adult literacy is associated with the use of preventive services, comprehension of medical conditions and adherence to medical instructions, self-management skills, physical and mental health, mortality, and health care costs.[1,12,13] Literacy is more strongly associated with these outcomes than educational attainment.[12,13] Whereas patients with limited literacy are more likely to be elderly, socioeconomically disadvantaged, live in rural areas, and belong to ethnic/minority groups, research has demonstrated an increased risk for poorer health with limited literacy beyond these characteristics, and that literacy may play a mediating role in health disparities.[14,15]

These findings underscore the need for health care interventions to address limited literacy, and a screening program might potentially aid physicians and other providers by identifying those at high risk. However, to justify a screening program, several conditions must be met: 1) Screening tests need to accurately and reliably detect limited literacy; 2) the benefit of early treatment options to reduce adverse health outcomes must be proven and available; and 3) the benefits need to outweigh adverse effects of the program.[11]

CURRENT SCREENING TESTS

Several screening instruments for literacy are available. These instruments have been well tolerated in research settings where they have been used extensively.[7–9,15] Yet it is uncertain whether patients would respond differently when tested by clinical staff with whom they have a relationship, and during times when they may be ill, anxious, and expecting medical care.

The most common literacy assessments in health literacy research include the Short version of the Test of Functional Health Literacy in Adults (S-TOFHLA) and Rapid Estimate of Adult Literacy in Medicine (REALM).[15] The S-TOFHLA takes 7–8 minutes to administer, assessing reading comprehension of

passages using the Cloze procedure. A 4-item numeracy section was originally included, but since removed, reducing the time to administer. As a medical word pronunciation test, the REALM requires less than 3 minutes to complete and is more widely used than the S-TOFHLA.

More recently, shorter tests have been developed with the notion they might be used as clinical screening tools.[7–9] An 8- and 7-word REALM are available and can be administered in less than 1 minute, yet their utility in research, let alone clinical settings, is less known.[7,8] Correlations with the Wide Reading Achievement Test for the 8- and 7-word REALM short forms were 0.64 and 0.83, respectively, and both exhibited moderate to high accuracy in classifying patients with low literacy skills (>85%). Weiss and colleagues recently developed another literacy screening instrument called The Newest Vital Sign (NVS), which assesses reading numeracy skills by testing understanding of information included on a nutrition label.[9] The NVS only requires 2–3 minutes to administer; however, in its current form, the NVS is highly sensitive and misclassifies patients with adequate literacy (area under the receiver operating characteristic (ROC) curve ranges from 0.71 to 0.88). One small study (N=119) found scores on the NVS not to be predictive of health outcomes, whereas S-TOFHLA scores were.[16] Possibly, the greater emphasis on health numeracy skills displayed in the NVS makes this a more complex set of tasks, which may limit its ability to differentiate patients compared to the S-TOFHLA and REALM.

In another approach, screening questions relating to literacy activities (e.g., "How confident are you filling out medical forms by yourself?") appear to be an acceptable and benign way to gauge literacy level, but do not perform as well as tests that evaluate reading (ROC=0.83) and perform no better than a prediction model of limited literacy based on demographic characteristics.[17]

POTENTIAL BENEFITS OF LITERACY SCREENING TO PATIENTS

The rationale for clinically assessing a patient's individual capacity for reading, understanding, and acting on health information is grounded in the premise that individuals with limited literacy have different communication and learning needs. This notion is supported by learning disabilities research, which classifies individuals with reading difficulties as either having developmental dyslexia or as persistently poor readers who face greater cognitive challenges.[18] In education, different paths for remediation have been proposed. In health care settings, specialized educational strategies based on individual learning requirements have not been established. As patient education would likely be more intensive for individuals with limited literacy, screening could potentially guide the allocation of resources.

To date, however, all suggested patient education and clinician–patient communication approaches for patients with limited literacy have been shown to benefit all patients and harm none.[6] Patients across all levels of literacy benefit from health materials that are easier to read.[19] Similarly, clinicians should learn how to communicate without jargon and confirm patient comprehension with all patients.[20] Whereas additional research evaluating these communication practices is warranted, it appears that practices such as confirming comprehension should not be reserved only for those with limited literacy, as clear health communication is not a scarce resource to be selectively distributed. There have been only 2 identified exceptions wherein patients with limited literacy received services necessary only to them, namely, adult basic education. In 1 pilot study, patients being treated for depression who were referred to an adult basic education program had lower levels of depressive symptoms compared to those not referred.[21] In another small study, parents of children in Head Start who participated in a supplementary Family Service Center program that included adult basic education had improved functional literacy scores, increased family incomes, and decreased depression in comparison with subjects who did not participate in this 2-year multimodal program.[22]

POTENTIAL FOR HARM

Sensitive and potentially stigmatizing topics are frequently broached in medical settings, justified by the overall likelihood of benefit from directed medical care that might result. In Seligman's study of literacy screening, 150 of the 160 (94%) subjects felt that health literacy screening was useful; however, as noted by the authors, without a direct measure of patient stigma, this study did not evaluate the possibility that patients may have nonetheless felt stigmatized.[10] Literacy screening programs could negatively impact patient care by promulgating fear and labeling. Previous research has found that nearly half of individuals with limited literacy report feelings of shame and often attempt to conceal this knowledge from others.[23,24] In 1 study, two-thirds of patients with limited literacy had never told their spouse and 19% had not disclosed their reading difficulties to anyone.[23] Thus, patients, especially those with lower literacy, may not be receptive to routine measurement of their literacy skills, nor want it recorded in their medical record.[24] Such shame could further alienate patients who already face a significant barrier accessing health care.

RECOMMENDATION

There are tools now available that, with further evaluation, could potentially be used in a literacy screening program. However, there is no known benefit from screening in the form of enrollment in a training program that would be delivered solely to those found to have limited literacy. While not definitive, there is also potential for harm, in the form of shame and alienation, which might be induced through clinical screening. At this time, there is insufficient evidence to recommend clinical screening for health literacy.

DISCUSSION

If interventions emerge that should be exclusively delivered to patients with limited literacy, there would be a more clear justification for a screening program. Additional research is warranted at this time to provide evidence of the utility of literacy tests in clinical settings, or possibly reveal screening techniques that truly minimize the risk of stigma and alienation. It should be noted that the trial described by Seligman and colleagues did not provide the physicians with specific

training, nor were they or their patients supported with additional patient education or disease management tools for those identified as having limited health literacy. Given that physicians in that trial discussed the result of literacy screening in only 2% of the encounters, a research agenda for clinical screening will likely need to include provider training and appropriate support materials to promote patient education and further reduce stigma.

Until research has shown a literacy screening program that can benefit patients without evidence of harm, physicians and other health professionals should pursue responses to the problem of limited health literacy that do not depend on screening. First, health plans and large provider groups should get an estimate of their local prevalence of limited literacy, by going to http://www.casas.org/lit/litcode/Search.cfm.[25] Rates reported on this website are derived from the 1993 National Adult Literacy Survey, and should be adjusted upwards for clinical populations that are older and have chronic diseases. Public health initiatives should provide or direct clinicians to this information. Along these lines, data from the 2004 National Assessment of Adult Literacy should be used to develop updated and more refined local estimates to inform providers.[26] Second, clinicians and health system administrators should work to reduce unneeded complexity. Discussion of screening focuses on patient skills, however, solutions ultimately must also be sought through streamlining an inordinately complicated health care system.[2] Third, "universal precautions" should be adopted to confirm all patients' understanding of critical self-care activities and to support problem solving.[27] Any significant patient education initiative will require dedicated resources. If clinicians take the time to evaluate patient comprehension, they can target each patient's specific clinical needs, rather than a specific group of patients.

Acknowledgment: *Dr. Wolf is supported by a career development award through the Centers for Disease Control and Prevention (1 K01 EH000067-01).*

Conflict of Interest: *None disclosed.*

Corresponding Author: *Michael K. Paasche-Orlow, MD, MA, MPH; Section of General Internal Medicine, Boston University School of Medicine, 91 East Concord Street, Suite 200, Boston, MA 02118, USA (e-mail: mpo@bu.edu).*

REFERENCES

1. **Dewalt DA, Berkman ND, Sheridan S, Lohr KN, Pignone MP.** Literacy and health outcomes: a systematic review of the literature. J Gen Intern Med. 2004;19:1228–39.
2. Institute of Medicine. Health literacy: A prescription to end confusion. In: **Nielsen-Bohlman L, Panzer A, Kindig DA, eds.** Washington, D.C., National Academies Press, 2004.
3. **Berkman N, Pignone MP, DeWalt D, Sheridan S.** Health literacy: Impact on Health Outcomes. Evidence Report/Technology Assessment. Rockville, MD:AHRQ, 2004.
4. Ad Hoc Committee on Health Literacy, American Medical Association Council on Scientific Affairs. Health literacy: report of the Council on Scientific Affairs. JAMA. 1999;281:552–7.
5. Joint Commission for the Accreditation of Hospital Organization (JCAHO). What Did the Doctor Say? Improving Health Literacy to Protect Patient Safety. Health Care at the Crossroads Report. Oakbrook, IL, 2007. February.
6. **Pignone M, DeWalt DA, Sheridan S, Berkman N, Lohr KN.** Interventions to improve health outcomes for patients with low literacy. A systematic review. J Gen Intern Med. 2005;20:185–92.
7. **Bass PF 3rd, Wilson JF, Griffith CH.** A shortened instrument for literacy screening. J Gen Intern Med. 2003;18:1036–8.
8. **Wolf MS, Heckinger EA, Arozullah A, et al.** Identifying persons with poor health literacy in 7 seconds: an important adjunct to improving cancer screening and treatment. J Clin Oncol 2004;22(14S):6007.
9. **Weiss BD, Mays MZ, Martz W, et al.** Quick assessment of literacy in primary care: the newest vital sign. Ann Fam Med. 2005;3:514–22.
10. **Seligman HK, Wang FF, Palacios JL, et al.** Physician notification of their diabetes patients' limited health literacy. A randomized, controlled trial. J Gen Intern Med. 2005;20:1001–7
11. United States Preventive Services Task Force. The guide to clinical preventive services. Rockville, MD. AHRQ, 2006.
12. **Wolf MS, Gazmararian JA, Baker DW.** Health literacy and functional health status among older adults. Arch Intern Med. 2005;165:1946–52.
13. **Wolf MS, Feinglass JM, Carrion V, Gazmararian J, Baker D.** Literacy and mortality among Medicare enrollees. J Gen Intern Med. 2006;21:81. Abstract.
14. **Wolf MS, Knight SJ, Durazo-Arvizu R, et al.** Associations of race, literacy, and income with PSA level among men with newly diagnosed prostate cancer. Urology. 2006;68:89–93.
15. **Paasche-Orlow MK, Parker RM, Gazmararian JA, et al.** The prevalence of limited health literacy. J Gen Intern Med. 2005;20(2):175–84.
16. **Osborn CY, Weiss BD, Davis TC, Skripkauskas S, Rodrigue C, Bass PF, Wolf MS.** Measuring adult literacy in health care: performance of the Newest Vital Sign. Am J Health Behav. 2007 Sep–Oct;31(Suppl 1):S36–46.
17. **Wallace LS, Rogers ES, Roskos SE, Holiday DB, Weiss BD.** Brief report: screening items to identify patients with limited health literacy skills. J Gen Intern Med. 2006;21:874–7.
18. **Shaywitz SE, Shaywitz BA.** Dyslexia (specific reading disability). Biol Psychology 2005;57:1301–1309.
19. **Weiss BD, Francis L, Senf JH, et al.** Literacy education as treatment for depression in patients with limited literacy and depression: a randomized controlled trial. J Gen Intern Med. 2006;21(8):823–8.
20. **Sudore RL, Landefeld CS, Williams BA, Barnes DE, Lindquist K, Schillinger D.** Use of a modified informed consent process among vulnerable patients: a descriptive study. J Gen Intern Med. 2006;21 (8):867–73.
21. **Paasche-Orlow MK, Riekert KA, Bilderback A, Chanmugam A, Hill P, Rand CS, Brancati FL, Krishnan JA.** Tailored education may reduce health literacy disparities in asthma self-management. Am J Respir Crit Care Med. 2005;172(8):980–6.
22. **Poresky RH, Daniels AM.** Two-year comparison of income, education, and depression among parents participating in regular Head Start or supplementary Family Service Center Services. Psychol Rep. 2001;88(3 Pt 1):787–96.
23. **Parikh NS, Parker RM, Nurss JR, Baker DW, Williams MV.** Shame and health literacy: the unspoken connection. Patient Educ Couns. 1996; 27:33–9.
24. **Wolf MS, Williams MV, Parker RM, Parikh NS, Nowlan AW, Baker DW.** Patients' shame and attitudes toward discussing the results of literacy screening. J Health Commun. 2007;12:1–12.
25. Comprehensive Adult Student Assessment System (CASAS). San Diego, CA. (Accessed March, 6, 2007, at http://www.casas.org/lit/litcode/Search.cfm.).
26. **Lurie N, Parker RM.** Moving health literacy from the individual to the community. Am J Health Behav. 2007;31:6–7.
27. **Paasche-Orlow MK, Schillinger D, Greene SM, Wagner EH.** How health care systems can begin to address the challenge of limited literacy. J Gen Intern Med. 2006;21(8):884–7.

Variation in Estrogen-Related Genes Associated with Cardiovascular Phenotypes and Circulating Estradiol, Testosterone, and Dehydroepiandrosterone Sulfate Levels

Inga Peter, Alyson Kelley-Hedgepeth, Caroline S. Fox, L. Adrienne Cupples, Gordon S. Huggins, David E. Housman, Richard H. Karas, Michael E. Mendelsohn, Daniel Levy, and Joanne M. Murabito

Center for Genetic Epidemiology and Modeling, Institute for Clinical Research and Health Policy Studies (I.P.), and Molecular Cardiology Research Institute (A.K.-H., G.S.H., R.H.K., M.E.M.), Tufts Medical Center, Boston, Massachusetts 02111; Center for Cancer Research (D.E.H.), Massachusetts Institute of Technology, Cambridge, Massachusetts 02139; National Heart, Lung, and Blood Institute's Framingham Heart Study (D.L., J.M.M., C.S.F.), Framingham, Massachusetts 01702; Biostatistics Department (L.A.C.), Boston University School of Public Health, Boston, Massachusetts 02115; Department of Preventive Medicine and Cardiology (D.L., J.M.M), Section of General Internal Medicine, Boston University School of Medicine, Boston, Massachusetts 02118; and the Center for Population Studies (D.L.), National Heart, Lung, and Blood Institute, Bethesda, Maryland 20892

Background: Younger age at the onset of menopause and lower circulating levels of estrogen are risk factors for cardiovascular disease. Several studies have detected associations between variations in genes encoding estrogen receptors α (ESR1) and β (ESR2), and enzyme aromatase (CYP19A1), which regulates the estrogen to testosterone ratio, and cardiovascular phenotypes in the Framingham Heart Study. To explore potential mechanisms by which these gene variants may contribute to cardiovascular disease, we tested the hypothesis that the polymorphisms were associated with endogenous steroid hormone levels.

Methods: Multiple regression analysis was used to assess the relation between reported polymorphisms and total serum estradiol, testosterone, and dehydroepiandrosterone sulfate levels in 834 men and 687 women who attended the third and fourth Framingham Heart Study examination cycles.

Results: In men, significant associations were detected between *CYP19A1* polymorphisms and estradiol and testosterone levels, and the estradiol to testosterone ratio (P ranges 0.0005–0.01). Specifically, carriers of common haplotype *rs700518*[*G*]-*(TTTA)*$_n$ [*L*]-*rs726547*[*C*] had higher estradiol levels (5% per copy; $P = 0.0004$), lower testosterone levels (17% per copy; $P = 0.036$), and a higher estradiol to testosterone ratio (24% per copy; $P < 0.0001$) compared with the *rs700518*[*A*]-*(TTTA)*$_n$ [*S*]-*rs726547*[*C*] carriers. In addition, postmenopausal carriers of the ESR2 *(CA)*$_n$ long allele and *rs1256031* [*C*] allele had moderately higher estradiol levels ($P \leq 0.03$). No significant associations with the ESR1 variants were detected.

Conclusions: Our findings suggest that variations in *CYP19A1* correlate with steroid hormone levels in men. Knowledge that a specific carrier status may predispose to altered steroid hormone levels may lead to targeted intervention strategies to reduce health risks in genetically susceptible individuals. **(*J Clin Endocrinol Metab* 93: 2779–2785, 2008)**

Male sex and the age of onset of menopause in women are independent factors that significantly increase the risk of hypertension, ventricular hypertrophy, and cardiac events, suggesting an important role for sex hormones in the etiology of cardiovascular disease (CVD) (1). A recent study has reported a lower risk of CVD events in older men with higher serum estra-

0021-972X/08/$15.00/0

Printed in U.S.A.

doi: 10.1210/jc.2008-0106 Received January 16, 2008. Accepted April 22, 2008.

First Published Online April 29, 2008

Abbreviations: BMI, Body mass index; CVD, cardiovascular disease; DHEAS, dehydroepiandrosterone sulfate; ER, estrogen receptor; ESR1, estrogen receptor α gene ESR2, estrogen receptor β gene; FHS, Framingham Heart Study; HWE, Hardy-Weinberg equilibrium; LD, linkage disequilibrium; SNP, single nucleotide polymorphism.

diol levels (2). Longer lifetime exposure to ovarian estrogens has protected against ischemic stroke (3). Moreover, low levels of dehydroepiandrosterone sulfate (DHEAS), the most abundant steroid mainly produced in the human adrenal gland and converted into potent androgens and/or estrogens through a series of enzymatic reactions, have been related to a greater risk of CVD in some (4, 5), but not in other, studies (2, 6, 7). However, a controversy exists regarding the CVD risk associated with serum testosterone concentrations (2, 4, 7).

Recently, genetic variations of the major proteins involved in steroid hormone conversion and receptor function have been described as significant contributors to cardiac disease susceptibility, attracting increased attention to genes implicated in estrogen metabolism. Specifically, numerous studies have found a significant association between polymorphisms in genes encoding estrogen receptors (ERs) α and β (*ESR1* and *ESR2*, respectively), which are expressed in the cardiovascular system, and myocardial infarction (8, 9), coronary artery disease (10–12), elevated blood pressure (13), altered lipoprotein levels (14, 15), adiposity (16), and left ventricular mass (17–19). In addition, variation in the gene encoding aromatase cytochrome 450 (*CYP19A1*), the enzyme that catalyzes the conversions of testosterone to estradiol and defines the estradiol to testosterone ratio, has been associated with blood pressure (13) and abdominal obesity (20). However, despite these and other studies, the molecular mechanisms that explain the relationship of estrogen-related genes with CVD risk are unclear.

Therefore, given moderate heritability ranging around 57–63% for testosterone (21, 22), 29–74% for DHEAS (23, 24), and about 25% for estradiol (22), and significant associations reported of single nucleotide polymorphisms (SNPs) in *ESR1*, *ESR2*, and *CYP19A1* with cardiovascular phenotypes in the Framingham Heart Study (FHS) (9, 13, 15, 16, 18, 19), we hypothesized that these relationships are mediated through circulating steroid hormone levels. We tested whether polymorphisms in these genes correlate with serum DHEAS, testosterone, and estradiol levels, and the estradiol to testosterone ratio in the FHS participants. This information may help shed light on the pathophysiology of CVD and its risk factors.

Subjects and Methods

Study sample

Participants in this study included unrelated individuals from the FHS's Offspring Cohort Study described in detail elsewhere (25). They attended periodic clinical examination 3 (1984–1987) and/or 4 (1987–1990), underwent medical and menopausal history, physical examination, and blood collection, including measurement of sex hormone levels, whereas a DNA sample and consent for genetic analysis were obtained at a later examination cycle (1996–1997). Subjects were excluded from the study if they were receiving hormone replacement therapy (n = 84) or had one ovary surgically removed (n = 45). The total sample included 1521 participants, of them 687 women. Women were considered postmenopausal if their periods stopped for 1 yr or more at examination, or both of their ovaries were surgically removed, and premenopausal otherwise, as defined by detailed self-report. For women whose menopausal status changed between the two examination cycles, only postmenopausal levels were used. Time since menopause was calculated by subtracting the reported age at onset of menopause from the chronological age at each examination.

Weight and height were measured at each examination. Body mass index (BMI) was calculated as the weight in kilograms divided by the square of the height in meters. Smoking status was defined by the self-reported number of cigarettes smoked per day in the year preceding the examination; alcohol consumption was recorded from the participant's report in ounces per week. Physical activity, determined by questionnaire, was represented as the weighted sum of the proportion of a typical day spent sleeping and performing sedentary, slight, moderate, or heavy physical activities (26). The following weights were used: sleep/rest, 1.0; sedentary, 1.1; light activity, 1.5; moderate activity, 2.4; and heavy activity, 5.0.

All subjects gave written informed consent. The FHS protocol is approved by the Boston University Medical Center Institutional Review Board.

Hormone assays

Steroid hormone levels were measured in serum samples using RIAs (Diagnostic Products Corp., Los Angeles, CA) with interassay coefficients of variation of 11% for total testosterone, 4% for total estradiol, and 11% for DHEAS. Testosterone was measured in men only.

SNP genotyping

Participants' genomic DNA was extracted from peripheral blood leukocytes using standard methods. Genotyping for the individual SNPs in *ESR1* (*rs2077647*, *rs2234693*, *rs9340799*, and *rs1801132*), *ESR2* (*rs1256031* and *rs1256059*), and *CYP19A1* (*rs700518* and *rs726547*) was performed as described previously (13). *ESR1* $(TA)_n$, *ESR1* $(CA)_n$, *ESR2* $(CA)_n$, and *CYP19A1* $(TTTA)_n$ repeat polymorphisms were genotyped using restriction fragment length analyses (Table 1 and supplemental Table 1, which is published as supplemental data on The Endocrine Society's Journals Online web site at http://jcem.endojournals.org). Genotyping was blinded to participant characteristics.

Statistical analysis

Observed genotype frequencies were compared with those expected under Hardy-Weinberg equilibrium (HWE) using a χ^2 test. Given multiple alleles observed for the repeat polymorphisms and their bimodal distribution, the genotype carrier status for each variant was coded using the median number of repeat sequence base pairs as a cutoff. Specifically, genotype *LL* was assigned if both alleles contained at least the median number of base pairs [$\geq$176 for *ESR1* $(TA)_n$; $\geq$ 277 for *ESR1* $(CA)_n$; $\geq$ 162 for *ESR2* $(CA)_n$, and $\geq$ 298 for *CYP19A1* $(TTTA)_n$]; *SS* was assigned if both alleles were "short" [$<$176 for *ESR1* $(TA)_n$; $<$ 277 for *ESR1* $(CA)_n$; $<$ 162 for *ESR2* $(CA)_n$, and $<$ 298 for *CYP19A1* $(TTTA)_n$], and *LS* if one allele was "long," and another one was "short."

To address skewed distribution of steroid hormone levels, logarithmic transformation was applied before analysis, and all results are reported as geometrical means. To account for significant sex-specific differences in steroid hormone levels, each variable was analyzed separately by gender and menopausal status. Multivariate linear regression analyses were performed to assess genetic associations with circulating estradiol and testosterone levels in men and postmenopausal women, and with DHEAS levels in men and premenopausal and postmenopausal women. All analyses were adjusted for age, weight, smoking status, alcohol consumption, and a number of years after the onset of menopause (in postmenopausal women). For individuals with steroid levels measured at both examinations, mean levels across the two examinations were used. In secondary analyses, risk factors were compared between genotype groups using ANOVA.

Pairwise linkage disequilibrium (LD) was evaluated using Lewontin's D' (27). Haplotypes were inferred by the expectation-maximization-algorithm. To account for allelic interaction, haplotypes were used as predictors in the regression models along with the aforementioned covariates.

The nominal threshold for statistical significance of all analyses was set at 0.05 and was not adjusted for multiple testing. Greater credibility

TABLE 1. Polymorphism characteristics

Gene	dbSNP rs no.	Location	Nucleotide substitution	MAF	*P* value for HWE
ESR1	$(TA)_n$	Promoter	$(TA)_n$	0.50[a]	0.76
6q25.1	rs2077647	Exon 1 (Ser10Ser)	T/C	0.45	0.10
	rs2234693	Intron 1	T/C	0.45	0.70
	rs9340799	Intron 1	A/G	0.36	0.83
	rs1801132	Exon 4 (Pro325Pro)	C/G	0.23	0.08
	$(CA)_n$	Intron 5	$(CA)_n$	0.36[a]	0.48
ESR2	rs1256059	Intron 2	C/T	0.44	0.77
14q23.2	$(CA)_n$	Intron 5	$(CA)_n$	0.21[a]	0.67
	rs1256031	Intron 7	T/C	0.48	0.58
CYP19A1	rs700518[b]	Exon 3 (Val80Val)	A/G	0.47	0.01
15q21.1	$(TTTA)_n$	Intron 4	$(TTTA)_n$	0.48[a]	0.99
	rs726547	Intron 4	C/T	0.05	0.78

dBSNP, The Single Nucleotide Polymorphism database (http://www.ncbi.nlm.nih.gov/(SNP/); MAF, minor allele frequency.

[a] Calculated using median number of repeats.

[b] In HWE in postmenopausal women.

was given to association results if a consistent trend was observed for SNPs in LD. All analyses were performed using SAS/STAT and SAS/Genetics software version 9.1 (SAS Institute Inc., Cary, NC).

Results

Nongenetic predictors of steroid hormone levels

The characteristics of the 834 male and 687 female unrelated eligible participants in the FHS Offspring cohort are shown in Table 2. Premenopausal women had the highest circulating estradiol levels, whereas postmenopausal women had the lowest estradiol and DHEAS levels. In multivariable-adjusted regression analyses, older age was a significant predictor of lower steroid hormone levels, except estradiol in men. Self-reported weekly alcohol intake was negatively correlated with DHEAS levels in both sexes and with estradiol in men, whereas smoking was associated with lower testosterone levels in men and DHEAS in premenopausal and postmenopausal women. Body weight was a better predictor of steroid hormone levels than BMI, and was positively correlated with estradiol and testosterone concentrations in men, but not women. The physical activity score was not found to be associated with DHEAS, estradiol, testosterone levels, and the estradiol to testosterone ratio in either men or premenopausal or postmenopausal women.

CYP19A1 association analysis

The genotype frequencies conformed to those expected by HWE, except *CYP19A1 rs700518*, which was in HWE in post-

TABLE 2. Participant characteristics by gender and menopausal status

Traits	Mean ± SD		
	Men (n = 834)	Premenopausal women (n = 347)	Postmenopausal women (n = 340)
Age (yr)	50.6 ± 9.7	42.6 ± 5.6	57.1 ± 6.4
BMI (kg/m^2)	27.6 ± 3.7	25.2 ± 5.4	26.5 ± 5.4
Weight (lb)	188.8 ± 28.7	148.2 ± 32.9	150.4 ± 31.8
Smoking (%)	26	29	26
Alcohol consumption (oz/wk)	4.2 ± 5.0	1.9 ± 3.0	1.9 ± 2.8
Physical activity (h/d)			
Sleep	7.3 ± 1.1	7.3 ± 1.1	7.3 ± 1.1
Sedentary activity	6.2 ± 3.1	6.4 ± 3.1	6.4 ± 3.1
Slight activity	5.9 ± 2.6	5.8 ± 2.6	5.7 ± 2.6
Moderate activity	3.5 ± 2.4	3.3 ± 2.4	3.3 ± 2.3
Heavy activity	1.1 ± 1.6	1.2 ± 1.8	1.3 ± 1.8
		Median (25th-75th percentile)	
Testosterone (ng/ml)[a]	5.6 (4.6–6.5)		
Estradiol (pg/ml)[a]	28.4 (21.9–37.3)		10.3 (3.0–10.3)
DHEAS (μg/dl)[a]	206 (136–327)	154 (106–220)	103 (70–155)
Years since menopause			10.2 ± 7.2

Measurements were available from both examinations for men (n = 606 for testosterone, n = 564 for estradiol, and n = 600 for DHEAS), premenopausal women (n = 533 for DHEAS), and postmenopausal women (n = 5 for estradiol and n = 254 for DHEAS).

[a] Values are raw examination measurements or averages over the two examinations.

menopausal women, but not in men and premenopausal women ($P = 0.01$). This polymorphism was in HWE in our previous study that tested a slightly different subset of the FHS unrelated individuals (13). Despite additional exclusion criteria applied to this study (*e.g.* availability of steroid hormone measurements, hormone replacement therapy, or surgical removal of one ovary), these factors are unlikely to select for or against this genetic variant, especially in men. Therefore, we assumed a random selection bias and included *CYP19A1 rs700518* in analysis.

In men, significant associations were detected between *CYP19A1 rs726547* and estradiol, testosterone, and the estradiol to testosterone ratio (P ranges between 0.01 and 0.0005) (Table 3). *CYP19A1* $(TTTA)_n$ repeat polymorphism as well as *CYP19A1 rs700518* were associated with estradiol concentrations ($P = 0.02$ and 0.005, respectively) and the estradiol to testosterone ratio ($P = 0.01$ and 0.006, respectively). Specifically, carriers of the minor alleles of *CYP19A1 rs700518* [*A*], $(TTTA)_n$ [*L*], and *rs726547* [*T*] alleles had higher estradiol, lower testosterone levels, and a higher estradiol to testosterone ratio than their noncarrier counterparts (Table 3).

CYP19A1 haplotype analysis

CYP19A1 rs700518, $(TTTA)_n$, and *rs726547* are in strong LD (pairwise D' ranges between 0.85 and 1.00), which resulted in three common haplotypes (frequency of $> 5\%$): H2, *rs700518* [*G*]- $(TTTA)_n$ [*L*]- *rs726547*[C] with the frequency of 45.3%; H7, *rs700518* [*A*]- $(TTTA)_n$ [*S*]- *rs726547*[*T*] with the frequency of 5.1%; and H8, *rs700518* [*A*]- $(TTTA)_n$ [*S*]- *rs726547*[C] with the frequency of 44.6%. Carriers of H2 had higher estradiol levels (1.16 pg/ml, or 5%, per each copy of haplotype; $P = 0.0004$), lower testosterone levels (0.94 pg/ml, or 17%, per each copy of haplotype; $P = 0.036$), and a higher estradiol to testosterone ratio (1.21, or 24%, per copy; $P < 0.0001$) compared with the H8 carriers.

TABLE 3. Adjusted circulating serum hormone levels in men by estrogen-related genotypes

				Estradiol		Testosterone		Estradiol to testosterone ratio	
Gene	SNP	Genotype	No.	Median (25th-75th percentile)	*P* value	Median (25th-75th percentile)	*P* value	Median (25th-75th percentile)	*P* value
CYP19A1	*rs700518*	G/G	210	26.73 (25.76–28.14)	0.005	5.51 (5.14–5.86)	0.60	4.86 (4.47–5.26)	0.006
		G/A	305	27.44 (26.29–28.83)		5.35 (5.05–5.73)		5.06 (4.74–5.60)	
		A/A	161	30.35 (29.42–31.65)		5.26 (4.99–5.65)		5.61 (5.33–5.96)	
	$(TTTA)_n$	S/S	190	26.32 (25.47–27.55)	0.02	5.45 (5.18–5.82)	0.51	4.78 (4.48–5.18)	0.01
		S/L	331	27.79 (26.82–29.10)		5.31 (5.04–5.63)		5.19 (4.87–5.61)	
		L/L	167	29.55 (28.62–30.60)		5.26 (5.01–5.67)		5.47 (5.19–5.78)	
	rs726547	C/C	713	27.22 (26.33–28.57)	0.03	5.40 (5.12–5.78)	0.01	5.01 (4.67–5.42)	0.0005
		C/T	85	29.89 (28.75–31.75)		5.14 (4.79–5.42)		5.87 (5.44–6.36)	
ESR1	$(TA)_n$	L/L	164	27.45 (26.69–28.71)	0.74	5.41 (5.16–5.73)	0.41	5.05 (4.73–5.36)	0.75
		L/S	338	28.34 (27.54–29.43)		5.34 (5.05–5.70)		5.25 (4.95–5.72)	
		S/S	164	27.53 (26.81–28.46)		5.28 (4.99–5.60)		5.09 (4.79–5.45)	
	rs2077647	T/T	224	27.57 (26.74–28.55)	0.41	5.26 (4.97–5.59)	0.22	5.16 (4.84–5.52)	0.21
		T/C	408	28.18 (27.31–29.49)		5.38 (5.06–5.75)		5.20 (4.88–5.65)	
		C/C	159	26.70 (25.75–27.98)		5.46 (5.22–5.87)		4.84 (4.52–5.22)	
	rs2234693	T/T	246	27.80 (26.96–28.96)	0.84	5.40 (5.08–5.76)	0.54	5.07 (4.78–5.45)	0.79
		T/C	400	27.50 (26.47–28.97)		5.32 (5.07–5.65)		5.15 (4.81–5.54)	
		C/C	160	27.18 (26.30–28.83)		5.47 (5.22–5.95)		4.97 (4.60–5.36)	
	rs9340799	A/A	327	27.99 (27.07–29.16)	0.81	5.32 (5.04–5.70)	0.59	5.18 (4.89–5.52)	0.69
		A/G	378	27.36 (26.40–28.90)		5.43 (5.13–5.75)		5.06 (4.72–5.48)	
		G/G	109	26.87 (26.19–28.50)		5.44 (5.21–5.94)		4.94 (4.51–5.32)	
	rs1801132	C/C	468	27.47 (26.47–28.81)	0.90	5.34 (5.07–5.68)	0.87	5.13 (4.78–5.48)	0.76
		C/G	269	27.58 (26.74–28.88)		5.41 (5.12–5.80)		5.08 (4.79–5.46)	
		G/G	45	27.85 (27.30–29.43)		5.39 (5.15–5.80)		5.05 (4.69–5.49)	
	$(CA)_n$	S/S	242	27.34 (26.60–28.55)	0.70	5.27 (4.94–5.59)	0.33	5.15 (4.79–5.58)	0.93
		S/L	297	28.26 (27.47–29.32)		5.44 (5.19–5.84)		5.11 (4.82–5.53)	
		L/L	79	28.01 (27.06–29.10)		5.23 (4.98–5.63)		5.27 (4.96–5.71)	
ESR2	*rs1256059*	C/C	225	28.06 (27.28–29.43)	0.56	5.32 (5.08–5.70)	0.18	5.23 (4.94–5.64)	0.20
		C/T	402	27.45 (26.53–28.59)		5.34 (5.02–5.66)		5.09 (4.76–5.51)	
		T/T	168	27.01 (26.15–28.17)		5.58 (5.31–6.01)		4.86 (4.52–5.22)	
	$(CA)_n$	S/S	515	27.22 (26.41–28.46)	0.20	5.42 (5.10–5.76)	0.49	5.01 (4.71–5.45)	0.60
		S/L	274	28.32 (27.38–29.73)		5.33 (5.08–5.69)		5.23 (4.95–5.58)	
		L/L	27	24.97 (24.13–25.71)		5.20 (4.95–5.58)		4.80 (4.55–5.02)	
	rs1256031	T/T	213	28.05 (27.24–29.59)	0.73	5.31 (5.07–5.69)	0.09	5.19 (4.89–5.56)	0.33
		T/C	403	27.37 (26.39–28.53)		5.33 (5.02–5.65)		5.13 (4.79–5.52)	
		C/C	194	27.43 (26.53–28.65)		5.61 (5.32–6.05)		4.90 (4.54–5.30)	

Adjusted for age, weight, smoking, and alcohol consumption.

ER gene association analysis

Postmenopausal women who carried the minor *ESR2* $(CA)_n$ [*L*] and *ESR2 rs1256031* [*C*] alleles, not in LD (D'= 0.005), had moderately higher estradiol levels ($P = 0.02$ and 0.03, respectively; Table 4). No significant associations of the steroid hormone phenotypes with the *ESR1* SNPs were detected.

Adjusted DHEAS means by genotype are shown in supplemental Tables 2 and 3.

Secondary analyses

Nongenetic predictors of steroid hormone levels by genotype are shown in supplemental Table 4. No consistent association of smoking status, alcohol consumption, and BMI with estrogen-related gene polymorphisms was found.

TABLE 4. Adjusted estradiol levels in postmenopausal women by estrogen-related genotypes

				Estradiol	
Gene	SNP	Genotype	No.	Median (25th-75th percentile)	*P* value
CYP19A1	*rs700518*	G/G	72	8.70 (5.75–13.83)	0.82
		G/A	138	10.19 (5.42–16.42)	
		A/A	72	9.88 (6.24–15.73)	
	$(TTTA)_n$	S/S	71	8.40 (5.36–13.94)	0.74
		S/L	154	9.24 (5.73–14.72)	
		L/L	64	10.15 (6.23–17.37)	
	rs726547	C/C	290	9.76 (6.04–14.98)	0.84
		C/T	33	10.11 (6.60–13.95)	
		T/T	1	13.20 (13.20–13.20)	
ESR1	$(TA)_n$	L/L	67	12.59 (7.09–21.07)	0.56
		L/S	149	9.26 (5.84–14.08)	
		S/S	73	11.04 (6.96–15.86)	
	rs2077647	T/T	92	12.10 (7.77–18.11)	0.26
		T/C	171	8.98 (5.79–14.95)	
		C/C	59	8.61 (4.92–13.86)	
	rs2234693	T/T	97	11.61 (7.64–17.57)	0.39
		T/C	172	9.00 (5.66–14.20)	
		C/C	60	11.62 (6.76–17.38)	
	rs9340799	A/A	131	11.04 (7.13–16.55)	0.62
		A/G	163	9.17 (5.73–14.10)	
		G/G	37	11.28 (6.31–13.97)	
	rs1801132	C/C	194	9.08 (5.58–14.58)	0.39
		C/G	101	12.29 (8.13–19.62)	
		G/G	21	9.51 (5.76–11.93)	
	$(CA)_n$	S/S	126	9.16 (5.54–13.26)	0.23
		S/L	115	10.96 (6.36–15.54)	
		L/L	40	16.28 (10.33–27.41)	
ESR2	*rs1256059*	C/C	102	8.33 (5.21–12.12)	0.22
		C/T	163	8.78 (5.49–13.95)	
		T/T	56	15.02 (10.84–24.61)	
	$(CA)_n$	S/S	213	10.70 (6.92–16.94)	0.03
		S/L	107	7.54 (4.55–11.56)	
		L/L	17	19.19 (13.06–21.01)	
	rs1256031	T/T	99	7.80 (4.92–10.88)	0.02
		T/C	158	8.72 (5.63–14.27)	
		C/C	74	15.73 (12.09–25.29)	

Adjusted for age, weight, smoking, and alcohol consumption.

Discussion

In this study we report association between variations in the aromatase gene (*CYP19A1*) and circulating estradiol and testosterone levels, as well as the estradiol to testosterone ratio, in men. It is well known that in postmenopausal women, when the ovaries cease to produce estrogens, and throughout life in men, estrogen is converted from steroid precursors through a series of enzymatic reactions. The enzyme aromatase is found in numerous tissues in the body where it catalyzes the conversion of testosterone into estrogens (Fig. 1). Experimental data have shown that disruption of *Cyp19* was associated with the development of obesity, decrease in lean mass, hypercholesterolemia, hyperleptinemia, and insulin resistance (reviewed in Ref. 28).

In men, peripheral aromatization of testosterone to active estrogen accounts for at least 75% of estrogen production (29), and the hormone balancing is defined by the testosterone to estrogen ratio. Age-associated testosterone decline has been shown in relation to diabetes, metabolic syndrome, reduced body lean mass (reviewed in Ref. 30), and mortality in men (31).

This study shows that common polymorphisms in *CYP19A1* are associated in a dose-response manner with higher circulating estradiol and lower testosterone levels, and, consequently, with a higher estradiol to testosterone ratio in men. The differences between homozygous carriers and noncarriers were 13% for estradiol levels, 5% for testosterone levels, and 17% for the estradiol to testosterone ratio. These differences are independent of genetic effects on BMI, smoking status, and alcohol consumption (supplemental Table 4). In men, lower testosterone levels and a reduced testosterone to estradiol ratio have been associated with coronary atherosclerosis (32), whereas testosterone insufficiency has been correlated with an increased risk of death over 20 yr, independent of risk factors and preexisting health conditions (33). Moreover, higher serum estradiol levels and an estradiol to testosterone ratio have been associated with a lower risk for CVD events in older men (2).

Our findings suggest that synonymous coding (*rs700518*)

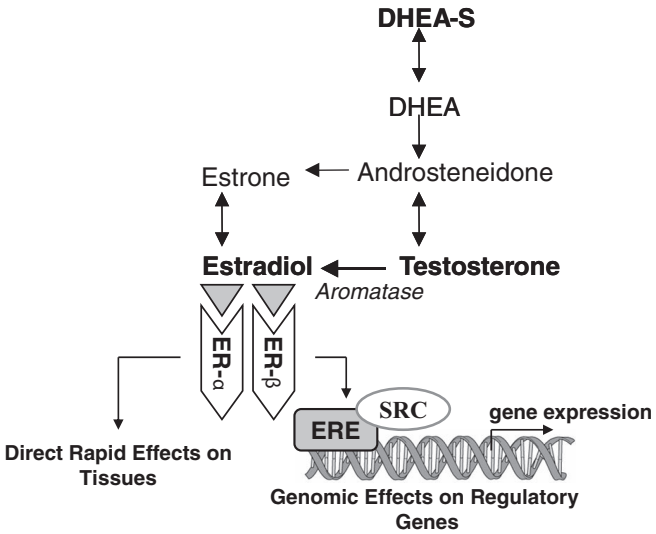


FIG. 1. Schematic summary of the role of the genes involved in estrogen metabolism. DHEA, Dehydroepiandrosterone; ERE, estrogen response element; SRC, steroid receptor coactivators.

and intronic (*rs726547*) nucleotide substitutions and the (*TTTA*)$_n$ repeat polymorphism in *CYP19A1* are linked to a gain-of-function variant or cause a splicing alteration that increases aromatase activity that results in the conversion of larger amounts of testosterone to active estrogen. In an earlier paper, a similar association between the "long" *CYP19A1* (*TTTA*)$_n$ allele and circulating estradiol has been reported in elderly men (34).

Estrogen exerts its actions directly by interacting with nuclear or membrane ERs, α and β. ERs are transcription factors that, when activated by estrogens, bind to estrogen-response elements in the promoter regions of target genes regulating their expression (Fig. 1). Variation in *ESR1* and *ESR2* may cause an impaired binding and, consequently, altered expression of genes regulating steroid hormone biosynthesis resulting in decreased circulating hormone levels. Therefore, we hypothesized that variations in the *ESR1* and *ESR2* genes, which have been repeatedly implicated in numerous cardiovascular phenotypes, decrease circulating estradiol, testosterone, and DHEAS levels, which, in turn, increase cardiovascular risks. Although we tested this hypothesis in the same cohort in which significant sex-specific associations were detected between the *ESR1* polymorphisms and higher risk of myocardial infarction (9), elevated blood pressure (13), altered lipoprotein particle size concentrations (14, 15), lower waist circumference (16), and more prominent age-related changes in left ventricular structure (18), as well as between higher serum estradiol levels and lower risk of CVD events (2), no relation of variation in *ESR1* with circulating steroid hormone levels was discovered. Nonetheless, a number of reports have shown significant associations between the two most studied variations in the first intron of *ESR1*, detected by digestion with restriction enzymes *PvuII*, *rs2234693*, and *XbaI*, *rs9340799*, and higher androstenedione, a precursor of testosterone (35), whereas conflicting results have been shown with serum levels of estradiol (36, 37), both in postmenopausal Caucasian women.

In addition, we detected moderate associations between postmenopausal estradiol levels and *ESR2 rs1256031* and *ESR2* (*CA*)$_n$. Carriers of minor alleles had higher circulating levels than their noncarrier counterparts. However, these findings were not supported by significant associations with another *ESR2* SNP, *rs1256059*, which is in tight LD with *ESR2 rs1256031*; yet, an association similar in direction and magnitude, though not statistically significant ($P = 0.22$), was detected (Table 4). Even though, our data do not help explain our previously reported finding that, after the adjustment for menopausal status, hypertensive *ESR2 rs1256031* [*TC*] female carriers had the largest left ventricular mass and wall thickness (19). Although we cannot completely exclude spurious associations common for this type of study, other reports have indicated significant relations of variation in *ESR2* with circulating estrogen and androgen levels in postmenopausal women (37, 38). Alternatively, we can speculate that the associations with the *ESR1* and *ESR2* SNPs, previously detected in the FHS, may be caused by hormone-independent effects of the receptors.

This study's limitations include the fact that active sex steroids are also synthesized locally in peripheral tissues, providing target tissues with controls that adjust the formation and metabolism of sex steroids to local requirements (39). These steroids are not released into the circulation and are not detectable in blood. However, it has been shown that circulating hormone levels reflect local steroid concentrations because their availability in the circulation ensures precursors for local synthesis (40). In addition, no formal adjustment for the multiple testing was performed. A standard correction for multiple hypothesis testing relies on the assumption that all statistical comparisons are independent. In the case of *CYP19A1*, three polymorphisms were in strong LD with one another, and steroid hormone levels were correlated. Nevertheless, some of our individual SNP and haplotype associations would independently pass the Bonferroni correction for multiple testing per gene ($\alpha = 0.05/21$ tests $= 0.002$ for *CYP19A1*), which represent a possible false-positive finding in one out of 2000 occurrences. The consistency and biological relevance of these associations should motivate further research to verify and extend these findings.

In summary, our findings suggest that associations between the *ESR1* gene polymorphisms and the incidence of myocardial infarction, elevated blood pressure, altered lipoprotein levels, and left ventricular structure are unlikely to be mediated through circulating steroid hormone levels. Importantly, the correlation of common genetic variations in *CYP19A1* with estradiol and testosterone concentrations and their ratio in men may have important health effects. Knowledge that a specific carrier status may predispose to altered steroid hormone levels, which can promote CVD risks, may lead to targeted intervention strategies to reduce health risks in genetically susceptible individuals.

Acknowledgments

Address all correspondence and requests for reprints to: Inga Peter, Ph.D., Assistant Professor of Medicine, Tufts Medical Center, 800 Washington Street, Box no. 63 Boston, Massachusetts 02111. E-mail: ipeter@tuftsmedicalcenter.org.

This work was supported by the National Heart, Lung and Blood Institute's project Grant HL077378 and training Grant HL069770 (to M.E.M.).

Disclosure Statement: The authors have nothing to disclose.

References

1. **Mendelsohn ME, Karas RH** 1999 The protective effects of estrogen on the cardiovascular system. N Engl J Med 340:1801–1811
2. **Arnlov J, Pencina MJ, Amin S, Nam BH, Benjamin EJ, Murabito JM, Wang TJ, Knapp PE, D'Agostino Sr RB, Bhasin S, Vasan RS** 2006 Endogenous sex hormones and cardiovascular disease incidence in men. Ann Intern Med 145: 176–184
3. **de Lecinana MA, Egido JA, Fernandez C, Martinez-Vila E, Santos S, Morales A, Martinez E, Pareja A, varez-Sabin J, Casado I** 2007 Risk of ischemic stroke and lifetime estrogen exposure. Neurology 68:33–38
4. **Mitchell LE, Sprecher DL, Borecki IB, Rice T, Laskarzewski PM, Rao DC** 1994 Evidence for an association between dehydroepiandrosterone sulfate and nonfatal, premature myocardial infarction in males. Circulation 89:89–93
5. **Feldman HA, Johannes CB, Araujo AB, Mohr BA, Longcope C, McKinlay JB** 2001 Low dehydroepiandrosterone and ischemic heart disease in middle-aged men: prospective results from the Massachusetts Male Aging Study. Am J Epidemiol 153:79–89
6. **Kiechl S, Willeit J, Bonora E, Schwarz S, Xu Q** 2000 No association between dehydroepiandrosterone sulfate and development of atherosclerosis in a pro-

spective population study (Bruneck Study). Arterioscler Thromb Vasc Biol 20:1094–1100
7. **Contoreggi CS, Blackman MR, Andres R, Muller DC, Lakatta EG, Fleg JL, Harman SM** 1990 Plasma levels of estradiol, testosterone, and DHEAS do not predict risk of coronary artery disease in men. J Androl 11:460–470
8. **Schuit SC, Oei HH, Witteman JC, Geurts van Kessel CH, van Meurs JB, Nijhuis RL, van Leeuwen JP, de Jong FH, Zillikens MC, Hofman A, Pols HA, Uitterlinden AG** 2004 Estrogen receptor α gene polymorphisms and risk of myocardial infarction. JAMA 291:2969–2977
9. **Shearman AM, Cupples LA, Demissie S, Peter I, Schmid CH, Karas RH, Mendelsohn ME, Housman DE, Levy D** 2003 Association between estrogen receptor alpha gene variation and cardiovascular disease. JAMA 290:2263–2270
10. **Lu H, Higashikata T, Inazu A, Nohara A, Yu W, Shimizu M, Mabuchi H** 2002 Association of estrogen receptor-α gene polymorphisms with coronary artery disease in patients with familial hypercholesterolemia. Arterioscler Thromb Vasc Biol 22:817–823
11. **Pollak A, Rokach A, Blumenfeld A, Rosen LJ, Resnik L, Dresner PR** 2004 Association of oestrogen receptor α gene polymorphism with the angiographic extent of coronary artery disease. Eur Heart J 25:240–245
12. **Rokach A, Pollak A, Rosen L, Friedlander Y, Blumenfeld A, Reznik L, Dresner-Pollak R** 2005 Estrogen receptor α gene polymorphisms are associated with the angiographic extent of coronary artery disease. J Clin Endocrinol Metab 90: 6556–6560
13. **Peter I, Shearman AM, Zucker DR, Schmid CH, Demissie S, Cupples LA, Larson MG, Vasan RS, D'Agostino RB, Karas RH, Mendelsohn ME, Housman DE, Levy D** 2005 Variation in estrogen-related genes and cross-sectional and longitudinal blood pressure in the Framingham Heart Study. J Hypertens 23:2193–2200
14. **Shearman AM, Demissie S, Cupples LA, Peter I, Schmid CH, Ordovas JM, Mendelsohn ME, Housman DE** 2005 Tobacco smoking, estrogen receptor α gene variation and small low density lipoprotein level. Hum Mol Genet 14: 2405–2413
15. **Demissie S, Cupples LA, Shearman AM, Gruenthal KM, Peter I, Schmid CH, Karas RH, Housman DE, Mendelsohn ME, Ordovas JM** 2006 Estrogen receptor-α variants are associated with lipoprotein size distribution and particle levels in women: the Framingham Heart Study. Atherosclerosis 185:210–218
16. **Fox CS, Yang Q, Cupples LA, Guo CY, Atwood LD, Murabito JM, Levy D, Mendelsohn ME, Housman DE, Shearman AM** 2005 Sex-specific association between estrogen receptor-α gene variation and measures of adiposity: the Framingham Heart Study. J Clin Endocrinol Metab 90:6257–6262
17. **Leibowitz D, Dresner-Pollak R, Dvir S, Rokach A, Reznik L, Pollak A** 2006 Association of an estrogen receptor-α gene polymorphism with left ventricular mass. Blood Press 15:45–50
18. **Peter I, Huggins GS, Shearman AM, Pollak A, Schmid CH, Cupples LA, Demissie S, Patten RD, Karas RH, Housman DE, Mendelsohn ME, Vasan RS, Benjamin EJ** 2007 Age-related changes in echocardiographic measurements: association with variation in the estrogen receptor-α gene. Hypertension 49: 1000–1006
19. **Peter I, Shearman AM, Vasan RS, Zucker DR, Schmid CH, Demissie S, Cupples LA, Kuvin JT, Karas RH, Mendelsohn ME, Housman DE, Benjamin EJ** 2005 Association of estrogen receptor β gene polymorphisms with left ventricular mass and wall thickness in women. Am J Hypertens 18:1388–1395
20. **Baghaei F, Rosmond R, Westberg L, Hellstrand M, Eriksson E, Holm G, Bjorntorp P** 2003 The CYP19 gene and associations with androgens and abdominal obesity in premenopausal women. Obes Res 11:578–585
21. **Meikle AW, Bishop DT, Stringham JD, West DW** 1986 Quantitating genetic and nongenetic factors that determine plasma sex steroid variation in normal male twins. Metabolism 35:1090–1095
22. **Ring HZ, Lessov CN, Reed T, Marcus R, Holloway L, Swan GE, Carmelli D** 2005 Heritability of plasma sex hormones and hormone binding globulin in adult male twins. J Clin Endocrinol Metab 90:3653–3658
23. **Rice T, Sprecher DL, Borecki IB, Mitchell LE, Laskarzewski PM, Rao DC** 1993 The Cincinnati Myocardial Infarction and Hormone Family Study: family resemblance for dehydroepiandrosterone sulfate in control and myocardial infarction families. Metabolism 42:1284–1290
24. **Nestler JE, Whitfield JB, Williams TY, Zhu G, Condon J, Kirk KM, Heath AC, Montgomery GW, Martin NG** 2002 Genetics of serum dehydroepiandrosterone sulfate and its relationship to insulin in a population-based cohort of twin subjects. J Clin Endocrinol Metab 87:682–686
25. **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
26. **Sherman SE, D'Agostino RB, Silbershatz H, Kannel WB** 1999 Comparison of past versus recent physical activity in the prevention of premature death and coronary artery disease. Am Heart J 138(5 Pt 1):900–907
27. **Lewontin RC** 1964 The interaction of selection and linkage. II. Optimum models. Genetics 50:757–782
28. **Murata Y, Robertson KM, Jones ME, Simpson ER** 2002 Effect of estrogen deficiency in the male: the ArKO mouse model. Mol Cell Endocrinol 193:7–12
29. **Hemsell DL, Grodin JM, Brenner PF, Siiteri PK, MacDonald PC** 1974 Plasma precursors of estrogen. II. Correlation of the extent of conversion of plasma androstenedione to estrone with age. J Clin Endocrinol Metab 38:476–479
30. **Kalyani RR, Dobs AS** 2007 Androgen deficiency, diabetes, and the metabolic syndrome in men. Curr Opin Endocrinol Diabetes Obes 14:226–234
31. **Shores MM, Matsumoto AM, Sloan KL, Kivlahan DR** 2006 Low serum testosterone and mortality in male veterans. Arch Intern Med 166:1660–1665
32. **Dunajska K, Milewicz A, Jedrzejuk D, Szymczak J, Kuliczkowski W, Salomon P, Bialy D, Poczatek K, Nowicki P** 2004 Plasma adiponectin concentration in relation to severity of coronary atherosclerosis and cardiovascular risk factors in middle-aged men. Endocrine 25:215–221
33. **Laughlin GA, Barrett-Connor E, Bergstrom J** 2008 Low serum testosterone and mortality in older men. J Clin Endocrinol Metab 93:68–75
34. **Gennari L, Masi L, Merlotti D, Picariello L, Falchetti A, Tanini A, Mavilia C, Del MF, Gonnelli S, Lucani B, Gennari C, Brandi ML** 2004 A polymorphic CYP19 TTTA repeat influences aromatase activity and estrogen levels in elderly men: effects on bone metabolism. J Clin Endocrinol Metab 89:2803–2810
35. **Zofkova I, Zajickova K, Hill M** 2002 The estrogen receptor α gene determines serum androstenedione levels in postmenopausal women. Steroids 67:815–819
36. **Schuit SC, de Jong FH, Stolk L, Koek WN, van Meurs JB, Schoofs MW, Zillikens MC, Hofman A, van Leeuwen JP, Pols HA, Uitterlinden AG** 2005 Estrogen receptor α gene polymorphisms are associated with estradiol levels in postmenopausal women. Eur J Endocrinol 153:327–334
37. **Sowers MR, Jannausch ML, McConnell DS, Kardia SR, Randolph Jr JF** 2006 Endogenous estradiol and its association with estrogen receptor gene polymorphisms. Am J Med 119(Suppl 1):S16–S22
38. **Westberg L, Baghaei F, Rosmond R, Hellstrand M, Landen M, Jansson M, Holm G, Bjorntorp P, Eriksson E** 2001 Polymorphisms of the androgen receptor gene and the estrogen receptor β gene are associated with androgen levels in women. J Clin Endocrinol Metab 86:2562–2568
39. **Labrie F, Luu-The V, Labrie C, Simard J** 2001 DHEA and its transformation into androgens and estrogens in peripheral target tissues: intracrinology. Front Neuroendocrinol 22:185–212
40. **Simpson E, Rubin G, Clyne C, Robertson K, O'Donnell L, Jones M, Davis S** 2000 The role of local estrogen biosynthesis in males and females. Trends Endocrinol Metab 11:184–188

2008; 30: 851–856 MEDICAL TEACHER

TWELVE TIPS

Twelve tips for excellent physical examination teaching

SUBHA RAMANI
Boston University School of Medicine, Boston, USA

Abstract

Background: Physical examination (PEx) skills are declining among medical trainees, yet many institutions are not teaching these systematically and effectively. Many variables contribute to effective teaching: teachers' confidence in their clinical skills, ability to demonstrate and assess these skills; availability of suitable patients; trainee attitude and fatigue; belief that institutions do not value clinical teachers. Finally, the relevance and significance of a systematic exam must be demonstrated or the teaching degenerates into a 'show-and-tell' exercise.

Aims: This paper describes twelve practical teaching tips that can be used to promote high quality PEx teaching in 5 minutes or 45 minutes.

Teaching tips: (1) Diagnostic hypotheses should guide reflective exam; (2) Teachers with the best clinical skills should be recruited; (3) A longitudinal and systematic curriculum can tailor teaching to multiple learner levels (4) Integration of simulation and bedside teaching can maximise learning; (5) Bedside detective work and games make learning fun; (6) The 6-step approach to teach procedures can be adopted to teach PEx; (7) Clinical teaching at the bedside should be increased; (8) Linking basic sciences to clinical findings will demonstrate relevance; (9) Since assessment drives learning, clinical skills should be systematically assessed; (10) Staff development can target improvement of teachers' clinical skills for effective teaching; (11) Technology should be used to study utility of clinical signs; (12) Institutions should elevate the importance of clinical skills teaching and recognize and reward teachers.

Conclusions: PEx is important in patient-physician interactions, a valuable contributor to accurate clinical diagnosis and can be taught effectively using practical tips. To reverse the trend of deficient clinical skills, precision of clinical findings should be studied and exam manoeuvres that do not contribute to diagnosis discarded; institutions should value clinical skills teaching, appoint and fund core faculty to teach and provide staff development to improve teaching skills.

Introduction

Teaching clinical skills in general and physical examination (PEx) in particular is unique and challenging compared to other methods of clinical teaching (Cox 1998). There are many variables that make up an effective PEx teaching encounter including: teachers' own clinical skills; trainees' prior knowledge, skills and interest; availability of patients with the necessary findings; patient willingness to be examined by a group of doctors and trainees who may not have any impact on their clinical care; the physical environment which is usually less than comfortable; trainee fatigue level etc. Even if all the factors mentioned above are perfectly aligned, the teachers themselves must be additionally skilled at demonstrating clinical signs and diagnosing when trainees have achieved those skills. Finally, the relevance of performing a systematic PEx and the significance of findings must be evident to trainees or the session quickly degenerates into a 'show-and-tell' exercise.

It has been noted that clinical skills teaching is not consistent across clinical rotations nor is it longitudinal. At our institution, for e.g., medical students learn PEx during two separate 'introduction to clinical medicine' blocks in their second year, sporadically and inconsistently during their third year clerkships, thereafter it is assumed that trainees have learned it all and not pursued systematically during their postgraduate years. There is marked variability in teaching across clerkships and the quality of the PEx training is entirely dependent on their clinical attachments and consultants. Often, it is their residents who end up teaching students and their own clinical skills are lukewarm at best.

Investigators suggest that a carefully obtained history and a focused physical exam contribute more to diagnoses than investigations alone even in the current medical environment (Peterson et al. 1992; Bordage 1995). To resuscitate clinical skills among clinicians, institutions need to raise the teaching of it to the highest priority, appoint core faculty to teach and invest in staff development.

Key challenges to effective teaching of physical examination

1. Teachers lack confidence in their own exam skills (Ramani et al. 2003).

Correspondence: Subha Ramani, MBBS, MMEd, MPH, Department of Medicine, Section of General Internal Medicine, Boston University School of Medicine, 715 Albany Street, Boston, Massachusetts-02118, USA. Tel: 617-638-7985; fax. 617-414-4676; email: sramani@bu.edu

ISSN 0142-159X print/ISSN 1466-187X online/08/09–00851–6 © 2008 Informa UK Ltd.
DOI: 10.1080/01421590802206747

2. Teachers lack the skills to demonstrate PEx effectively and efficiently (Cox 1998).
3. Medical schools teach the long winded systematic approach to exam without teaching history based hypotheses formulation and hypotheses based focused physical exam (Benbassat et al. 2005).
4. Lack of a systematic, longitudinal curriculum with different levels of teaching for different levels of trainees (Goldstein et al. 2005).
5. Lack of clarity in setting learning objectives for the teaching session, i.e. describing clearly what to look for and how to look for it (Cox 1998).
6. Not explaining what the findings are caused by, the differential diagnoses and clinical significance of findings (Cox 1998).
7. Teachers lecturing too much or demonstrating without providing trainees ample time and opportunities to practise (Cox 1998).
8. Deficient assessment of physical exam. This includes exclusive use of standardized patient based OSCE without observation in clinical practice as well as absence of physical exam assessment of postgraduate trainees in the United States (Holmboe 2004; Hatala et al. 2007b).
9. Medical institutions not valuing clinical skills enough and relying on technology to provide all the answers (Ramani et al. 2003).

The following twelve tips are practical strategies that institutions and individual teachers can use to promote high quality physical exam teaching (Table 1).

Tip 1

Physical exam should be reflective guided by diagnostic hypotheses

The traditional textbooks of clinical diagnosis describe very detailed and systematic physical exam of each organ system, which may be impractical in clinical settings where time is of the essence. Benbassat and colleagues (Benbassat et al. 2005) question the value of such a systematic head to toe exam and state that this method of exam divorces data collection from clinical reasoning. They suggest that challenging students to seek specific physical findings may increase the likelihood of detecting findings and may transform patient history and physical from routine activities into intellectually exciting experiences. A reflective examination is guided by diagnostic hypotheses, detecting cues for diagnosis in patients based on their complaints and formulating differential diagnosis.

A classic example is the neurologic exam which has traditionally been a time intensive exam. Some authors suggest classifying neurologic exam skills into essential, marginal and reserve skills (Glick 2005). Essential skills are mandatory steps and contribute most to diagnosis, marginal skills are steps of the exam that do not necessarily add to diagnostic hypotheses and reserve skills are those that would be used in specific situations to guide diagnosis. Glick argues that an evidence-based, progressive approach has the potential to impact positively teaching and efficiency of practice as well as externally mandated quality standards. This classification of PEx skills can be readily applied to other organ specific exam.

Tip 2

The most skilled clinicians should be recruited for PEx teaching

It appears that many institutions delegate junior faculty or senior trainees to teach PEx sessions. With the reported decline in clinical skills among trainees and practicing clinicians (Kern et al. 1985; Mangione et al. 1995; Mangione & Nieman 1999; Mangione 2001; Vukanovic-Criley et al. 2006), we might be allowing the blind to lead the blind. Allowing those whose clinical skills are not of the highest level to teach junior trainees will perpetuate bad technique, poor recognition of abnormal findings, inability to interpret their significance leading to lack of application in daily patient care. Institutions should actively recruit recently retired clinicians with excellent clinical examination skills to teach trainees, many of whom have trained before the era of obsession with high technology medicine.

The University of Washington underwent a major curriculum reform for teaching its students clinical skills in response to problems identified during needs assessment as well as problems reported about clinical skills of trainees nationwide (Goldstein et al. 2005). They established the college system with a core group of clinical teachers who committed a substantial portion of their time to teach and mentor medical students, as well as develop and assess the curriculum. They developed explicit benchmarks for teaching core clinical skills. They identified a core of expert teachers, ensured

Table 1. Suggested recommendations to enhance physical examination teaching.

Institutions	Teachers	Content
The best for the job	Acquire bedside teaching skills	Integrate basic sciences into physical exam teaching
Structured curriculum	Improve demonstration skills	Incorporate evidence based physical exam
Staff development	Reinforce and update own physical exam skills	Encourage use of technology to corroborate exam findings
Assessment		
Elevate importance of physical exam in trainee education	Use multiple teaching methods and make it fun	
Clinical skills centres		

adequate financial support for their teaching activities, provided them with a teaching and mentoring role, gave them a key role in curriculum development, and created an environment that refocused the school's resources on teaching and evaluating core clinical skills. Their preliminary data indicate that their program enjoys remarkable success with their students and faculty alike.

Tip 3

Establish a structured and longitudinal curriculum

Many professional societies have argued for the importance of teaching and evaluating clinical skills. But unpredictable patient exposure based on curricular needs may result in scant structured learning opportunities to acquire critical skills. Moreover, there are a variety of learning styles that different trainees use predominantly to learn. Therefore, a 'one size fits all' or a single structured type of educational approach would not enhance trainee ability to select and use specific history or physical exam skills for individual problems. To assure that essential physical exam skills are acquired, it most likely requires that both systematic instructional strategies (didactic and hands on) and repeated learning opportunities are available to reinforce learning (Allen et al. 1991).

Goldstein et al. (2005) at the University of Washington, felt it was necessary to have a 4-year integrated clinical skills curriculum and move away from the block based non-longitudinal traditional teaching of physical exam. In their curriculum, they delineated clinical competency domains for increasingly advanced levels with benchmarks clearly stated for each level of training. Such a longitudinal curriculum would also ensure that trainees continue to continuously learn clinical skills from basic to increasingly advanced levels building on what was previously learned. In the separated block curricula, students first learn normal exam, a few months later they are required to do a full history and exam on a patient most of which is not observed. During the third year clerkships, some attachments organize physical diagnosis sessions, others don't. This method does not enable recall or reinforcement of previously acquired skills and several assumptions of trainees' skills levels are held by faculty.

Issenberg & McGaghie (2002) state that a spiral curriculum1 is a useful approach to teach PE for the following reasons:

1. There is an iterative revisiting of core clinical skills throughout the entire medical school curriculum;
2. Clinical skills are revisited at numerous levels of difficulty;
3. Clinical skills are related to previous skills;
4. The competence and self-efficacy of students increases with each visit to a skill.

Tip 4

Integrate simulation with bedside learning

Using simulation can enable tutors to structure new learning opportunities, provide standardized and reproducible experiences and create learner centred environments where mistakes are permissible (Dent 2001). Clinical skills centres also provide ample opportunities to assess teachers' clinical and teaching skills as well as trainees' physical exam skills. Instead of using clinical skills centres to teach systematic exam skills without reinforcing this in clinical practice, Kneebone argues that the 2 environments should be a continuum (Kneebone et al. 2004). He proposes a more integrated approach that bridges the 2 learning environments. Learners should be able to practice in a simulated environment to prepare them for experiences with real patients and also return to the simulated environment when they had identified a clinical need from their experiences with real patients. They can reinforce what was learned at the bedside by listening to abnormal heart and lung sounds etc using a simulator. Repetition would reinforce the learning and enable them to better detect and recognise abnormal findings in future patients. They would be able to go at their own pace, with sustained practice and feedback such that learning could be immediately applied to the real situation in the workplace.

Tip 5

Teach it well: the 6 step approach

Educators from Denmark (Faarvang & Ringsted 2006) took the traditional 4-step approach to teaching procedural skills and expanded it to a 6-step approach to teach joint examination at their institution. These steps include the following and steps 0 and 5 were considered essential to effective teaching of physical exam:

0. assessing learners' needs and diagnosing their skill level;
1. the tutor demonstrates the procedure;
2. the tutor repeats the procedure and explains what is done;
3. the student instructs the tutor while he or she repeats the procedure;
4. the student demonstrates the procedure and explain what they do;
5. after observing learners in step 4, the teachers give feedback.

This principle is not specific for joint examination and can easily be applied to teaching examination skills for other groups of chronic patients with specific physical signs and symptoms. This approach should be included in institutional efforts at developing skilled teachers of clinical skills.

Tip 6

Make it fun – The clinical detectives

Most of the other tips relate to the importance and relevance of teaching PEx in medical education today as well as educational strategies for effective teaching. In addition, teachers should reflect on methods to make the teaching fun. Methods to teach physical diagnosis include bedside rounds, advanced physical diagnosis courses, utilizing senior trainees as teachers,

evidence- based physical diagnosis, use of teaching OSCEs and web-based curricula. Organising a game like atmosphere such as quizzes, medical jeopardy, physical findings treasure hunt, making the session an exercise in medical detection, making predictions about results of investigations based on exam findings can make PEx teaching more interactive, more engaging and more fun. At our institution, we have implemented more systematic PEx teaching for our residents this year. We have used a variety of teaching methods to keep the interest level high. Such methods include twice a month bedside morning reports, a marked change from the sit down case based discussions; physical exam jeopardy sessions with participants divided into teams and prizes for the winning team; interactive case based lectures focusing on PEx; and clinical problem solving sessions focusing on clinical exam. All of the sessions have been very interactive and received well by our residents.

Tip 7

Increase and improve bedside clinical teaching

Bedside teaching provides the best forum for clinical teachers to demonstrate physical examination techniques and teach physical exam. Yet, the frequency of bedside teaching is reported to have decreased from an incidence of 75% in the 1960s to less than 16% in the 1990s (Shankel & Mazzaferri 1986; Ende 1997). PEx skills cannot be taught in a classroom and require the presence of a patient, real or simulated. Although many clinical teachers find this an intimidating mode of teaching that bares their own deficiencies, they need to realize that all of them possess a wide range of clinical skills that they can teach their junior and far less experienced trainees (Ramani et al. 2003). This is particularly true if the patients are purely the subject of a physical exam teaching exercise and participants of the exercise are not members of the clinical team providing patient care. Some common sense strategies combined with faculty development programs at individual institutions can overcome some of this insecurity and promote bedside rounds which can be educational and fun for teachers and learners alike. Many strategies have been recommended in literature by educators including a twelve tips article (Ramani 2003) and three-domain model based on teacher-patient and team interactions (Janicik & Fletcher 2003). The twelve tips article simplifies key strategies for effective bedside teaching and has categorized teaching behaviours as those that can be carried out sequentially before rounds, during rounds and after rounds.

Tip 8

Technology can help not hinder physical exam teaching

Trainees and faculty may feel that the physical examination is a subjective art when compared with the more objective laboratory tests and imaging information (Andersen et al. 2001). A possible solution is using technology to aid the teaching of PEx. Physical diagnosis teachers today also have access to a wide variety of technological aids to enhance their teaching such as websites, CD-ROMs, simulators, videotapes, and infrared multiuser stethoscopes. Although technology has been often been described as a barrier to physical exam teaching, technology can actually be used to improve physical exam teaching, study the accuracy of physical signs and help clinical teachers discard those signs that have no proven value in diagnostic accuracy. The predictive value and clinical utility of many physical examination techniques and physical findings have been questioned and evaluated as illustrated in the the Rational Clinical Exam series in the JAMA journal. Continuing to study these issues will keep the teacher of physical diagnosis up to date and direct physical diagnosis teaching and evaluation to the clinically most useful techniques and findings (Ende & Fosnocht 2002).

Tip 9

Apply basic sciences to clinical exam

In their junior years, students prefer to see the relevance of the skills they are learning to what they learned in basic science, and how it will serve them in practice. In their senior years, students must continue to apply basic science to clinical medicine (Issenberg & McGaghie 2002). Physical examination, which is considered by many to be a cornerstone for clinical diagnosis relies on a solid foundation in basic sciences such as anatomy, physiology and pathology (AACA 2001). Where to inspect, percuss, palpate, and auscultate, as well as the proper analysis of physical findings during a physical examination, demand an anatomical foundation rooted in the following concepts: anatomical terminology, normal variation, three-dimensional relationships, functional and living anatomy, and most importantly, clinical anatomy (AACA 2001). Palpable bony landmarks provide important references for predicting underlying anatomy, eg the various valvular areas of the heart. Bones and bony landmarks are often used to define imaginary lines that serve as helpful references for estimating the location of underlying anatomical structures, eg the midclavicular line. Knowledge of neuroanatomy is essential to localise neurologic lesions based on exam and without knowledge of physiology, trainees cannot comprehend the mechanism behind the audibility of cardiac sounds and murmurs. Finally, knowledge of sciences such as pathology, immunology and microbiology is key in formulating differential diagnoses on conclusion of a physical exam.

Tip 10

Assessment drives curriculum

The use of standardized or simulated patients has, in many centres, become synonymous with the OSCE as an approach to the assessment of clinical competence. Little attention has been paid to when it is appropriate to use real patients (RPs) and when standardized simulated patients (SPs) should be used. Although RPs with stable clinical findings may be used,

in most assessments SPs without physical findings are employed. However, some medical educators have reported that the correlation between exam technique and diagnostic accuracy is poor (Szauter & Ainsworth 2006, Hatala et al. 2007a,b) and using SPs without physical findings may result in incomplete assessment of a trainee. Some factors needed to be attended to in planning assessment of clinical competence (Collins & Harden 1998):

1. what is being assessed, including the level of abnormality and level of interaction with the patient required;
2. the level of standardization required, with greater emphasis on standardization needed for high-stakes national examinations;
3. the logistics, including the availability and costs of real patients and trained simulated patients;
4. the context, for example, practice-based or formal examinations of the OSCE type;
5. the level of realism or authenticity required.

Thus, if detection of abnormal findings and diagnostic accuracy are to be assessed, real patients are needed and this is true for advanced level trainees whose exam technique alone cannot be used as a surrogate for accurate bedside diagnoses.

Another method that can be used to assess clinical skills in practice is the mini-CEX which consists of short bedside interactions of trainees observed and graded by faculty (Norcini et al. 2003). These can be formative and summative. Unfortunately, the quality of any assessment is dependent on faculty skills in history and physical exam, demonstration, observation, assessment and feedback and this has been a subject of some concern (Holmboe 2004). For all these reasons, staff development is crucial and institutions need to make an investment in the selection and training of their core clinical faculty as has been done very well by the University of Washington (Goldstein et al. 2005)

Tip 11

Staff development: Teach the teachers clinical skills

It has been said that 'To advance the art and science in clinical examination, the equipment a clinician most needs to improve is himself.'

Several investigators have reported that the clinical skills of trainees has declined since the 1970s (Kern et al. 1985; Mangione et al. 1995; Mangione 1999, 2001; Vukanovic-Criley et al. 2006), it becomes evident that these generations of trainees now turned faculty are teaching newer generations. Further, increasing pressures on faculty and curricular time may be leading to decreased attention to the teaching of the physical examination as well as decreased opportunity for faculty to improve their own skills. Cox states that since a major task in clinical teaching is to help students collect clinical evidence, teachers must be skilled in demonstrating how to do it (Cox 1998). To become skilful, each teacher must practise demonstration. To improve their skills teachers need to be observed during their teaching and provided with feedback on their effectiveness. And to be confident of the outcomes of effective demonstration, teachers must be sure that trainees have learned.

Providing periodic faculty development sessions to enhance their bedside teaching skills and to update them on new information about the utility (or lack of utility) of specific physical examination manoeuvres and findings would be very useful (Anderson et al. 2001). Residency program directors, chief residents, and other key clinical teachers should be encouraged to participate. Developing resident skill in physical diagnosis is critical, as they are the principal teachers of medical students and they need to be able to model these skills for students.

Tip 12

Integrate into the institutional values

Clinical teachers often perceive that in the current culture of academic medical institutions, teaching is an activity without commensurate rewards, financial or non-financial, and teaching is often the ball that is dropped before clinical work or research. Overcoming the cultural barriers to teaching requires a department or institution-wide initiative (Anderson et al. 2001). Although this cannot be expected to occur overnight, a faculty leader can promote innovative solutions with the support of the department chair, the medical school dean's office, and residency or clerkship directors. Medical schools should reward the teacher of physical diagnosis by acknowledging teaching efforts through personal recognition awards, salary as well as time allocated to enable faculty to teach.

They should also provide protected time and administrative support to the directors of physical exam courses. Faculty should be given time to teach by commensurate reduction of their clinical duties during the course. They should be rewarded for physical diagnosis teaching by including this activity in consideration of promotion decisions. Institutions could create awards that specifically recognize faculty or residents with outstanding physical diagnosis skills and make available funds for purchasing physical diagnosis teaching materials such as CD-ROMs, simulators, audiotapes, and videotapes. Finally, faculty with an interest in research in physical diagnosis should be supported and encouraged with start-up funds.

Conclusion

Physical examination is a key aspect of patient-physician interactions and a valuable contributor to accurate clinical diagnosis. There are many innovative ways to teach PE and it can be taught in a time efficient manner. Rather than being a barrier to improving clinical skills, technology can actually help establish accuracy of physical exam. The medical community should continue to study the precision of clinical findings and discard those exam manoeuvres that do not

contribute to patient diagnosis. Institutions should elevate the value of clinical skills teaching, appoint and fund core faculty to teach and provide staff development to improve both faculty physical exam skills as well as their PEx teaching skills. Diligent attention to teaching PEx exam may reverse the trend of deficient clinical skills among generations of medical trainees and put the patient back into patient care.

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

Notes on contributor

SUBHA RAMANI, MBBS, MMEd, MPH. Dr. Ramani is a general internist and medical educator. She is an Associate Professor of Medicine at the Boston University School of Medicine and directs Clinical Skills Development for the Internal Medicine Residency Program. She has completed a Masters in Medical Education at the University of Dundee. Her major areas of interest in medical education include clinical skills teaching and assessment, clinical teaching methods and staff development in teaching.

References

AACA. 2001. Clinical anatomy and the physical examination part I: Thorax, abdomen, perineum, and pelvis. Clin Anat 14(5):332–348.

Allen SS, Bland CJ, Harris IB, Anderson D, Poland G, Satran L, Miller W. 1991. Structured clinical teaching strategy. Med Teach 13(2):177–184.

Anderson RC, Fagan MJ, Sebastian J. 2001. Teaching students the art and science of physical diagnosis. Am J Med 110(5):419–423.

Benbassat J, Baumal R, Heyman SN, Brezis M. 2005. Viewpoint: Suggestions for a shift in teaching clinical skills to medical students: The reflective clinical examination. Acad Med 80(12):1121–1126.

Bordage G. 1995. Where are the history and the physical? Cmaj 152(10):1595–1598.

Collins JP, Harden RM. 1998. AMEE Medical Education Guide No. 13: Real patients, simulated patients and simulators in clinical examinations. Med Teach 20(6):508–521.

Cox K. 1998. How well do you demonstrate physical signs? Med Teach 20(1):6–9.

Dent J. 2001. Current trends and future implications in the developing of clinical skills centres. Med Teach 23(5):483–489.

Ende J. 1997. What if Osler were one of us? Inpatient teaching today. J Gen Intern Med 12(Suppl 2):S41–48.

Ende J, Posnocht KM. 2002. Clinical examination: Still a tool for our times? Trans Am Clin Climatol Assoc 113:137–150.

Faarvang KL, Ringsted C. 2006. A six-step approach to teaching physical examination. Med Educ 40:459–489.

Glick TH. 2005. Toward a more efficient and effective neurologic examination for the 21st century. Eur J Neurol 12:994–997.

Goldstein EA, Maclaren CF, Smith S, Mengert TJ, Maestas RR, Foy HM, Wenrich MD, Ramsey PG. 2005. Promoting Fundamental Clinical Skills: A Competency-Based College Approach at the University of Washington. Academic Medicine 80:423–433.

Hatala R, Cole G, Kassen BO, Bacchus CM, Issenberg SB. 2007a. Does physical examination competence correlate with bedside diagnostic acumen? An observational study. Med Teach 29(2-3):199–205.

Hatala R, Issenberg SB, Kassen BO, Cole G, Bacchus CM, Scalese RJ. 2007b. Assessing the relationship between cardiac physical examination technique and accurate bedside diagnosis during an objective structured clinical examination (OSCE). Acad Med 82(10 Suppl):S26–29.

Holmboe ES. 2004. Faculty and the observation of trainees' clinical skills. Problems and opportunities. Acad Med 79(1):16–22.

Issenberg SB, Mcgaghie WC. 2002. Clinical skills training – practice makes perfect. Med Educ 36:210–211.

Janicik RW, Fletcher KE. 2003. Teaching at the beside: A new model. Med Teach 25(2):127–130.

Kern DC, Patrino TA, Korst DR. 1985. The lasting value of clinical skills. JAMA 254(1):70–76.

Kneebone RL, Scott W, Darzi A, Horrocks M. 2004. Simulation and clinical practice: Strengthening the relationship. Med Educ 38(10):1095–1102.

Mangione S. 1999. The teaching of chest auscultation in U.S. internal medicine and family practice medicine residencies. Acad Med 74(10 Suppl):S90–92.

Mangione S. 2001. Cardiac auscultatory skills of physicians-in-training: A comparison of three English-speaking countries. Am J Med 110(3):210–216.

Mangione S, Burdick WP, Peitzman SJ. 1995. Physical diagnosis skills of physicians in training: A focused assessment. Acad Emerg Med 2(7):622–629.

Mangione S, Nieman LZ. 1999. Pulmonary auscultatory skills during training in internal medicine and family practice. Am J Respir Crit Care Med 159(4 Pt 1):1119–1124.

Norcini JJ, Blank LL, Duffy FD, Fortna GS. 2003. The mini-CEX: A method for assessing clinical skills. Ann Intern Med 138(6):476–481.

Peterson MC, Holbrook JH, Von Hales D, Smith NL, Staker LV. 1992. Contributions of the history, physical examination, and laboratory investigation in making medical diagnoses. West J Med 156(2):163–165.

Ramani S. 2003. Twelve tips to improve bedside teaching. Med Teach 25(2):112–115.

Ramani S, Orlander JD, Strunin L, Barber TW. 2003. Whither bedside teaching? A focus-group study of clinical teachers. Acad Med 78(4):384–390.

Shankel SW, Mazzaferri EL. 1986. Teaching the resident in internal medicine. Present practices and suggestions for the future. JAMA 256(6):725–729.

Szauter K, Ainsworth M. 2006. Use of standardized patients with actual findings. Med Educ 40:1143–1144.

Vukanovic-Criley JM, Criley S, Warde CM, Boker JR, Guevara-Matheus L, Churchill WH, Nelson WP, Criley JM. 2006. Competency in cardiac examination skills in medical students, trainees, physicians, and faculty: A multicenter study. Arch Intern Med 166(6):610–616.

2008; 30: 347–364 MEDICAL TEACHER

AMEE GUIDE

AMEE Guide no. 34: Teaching in the clinical environment

SUBHA RAMANI[1] & SAM LEINSTER[2]

[1]Boston University School of Medicine, USA, [2]University of East Anglia, Norwich, UK

Abstract

Teaching in the clinical environment is a demanding, complex and often frustrating task, a task many clinicians assume without adequate preparation or orientation. Twelve roles have previously been described for medical teachers, grouped into six major tasks: (1) the information provider; (2) the role model; (3) the facilitator; (4) the assessor; (5) the curriculum and course planner; and (6) the resource material creator (Harden & Crosby 2000).

It is clear that many of these roles require a teacher to be more than a medical expert. In a pure educational setting, teachers may have limited roles, but the clinical teacher often plays many roles simultaneously, switching from one role to another during the same encounter. The large majority of clinical teachers around the world have received rigorous training in medical knowledge and skills but little to none in teaching. As physicians become ever busier in their own clinical practice, being effective teachers becomes more challenging in the context of expanding clinical responsibilities and shrinking time for teaching (Prideaux et al. 2000). Clinicians on the frontline are often unaware of educational mandates from licensing and accreditation bodies as well as medical schools and postgraduate training programmes and this has major implications for staff training. Institutions need to provide necessary orientation and training for their clinical teachers. This Guide looks at the many challenges for teachers in the clinical environment, application of relevant educational theories to the clinical context and practical teaching tips for clinical teachers. This guide will concentrate on the hospital setting as teaching within the community is the subject of another AMEE guide.

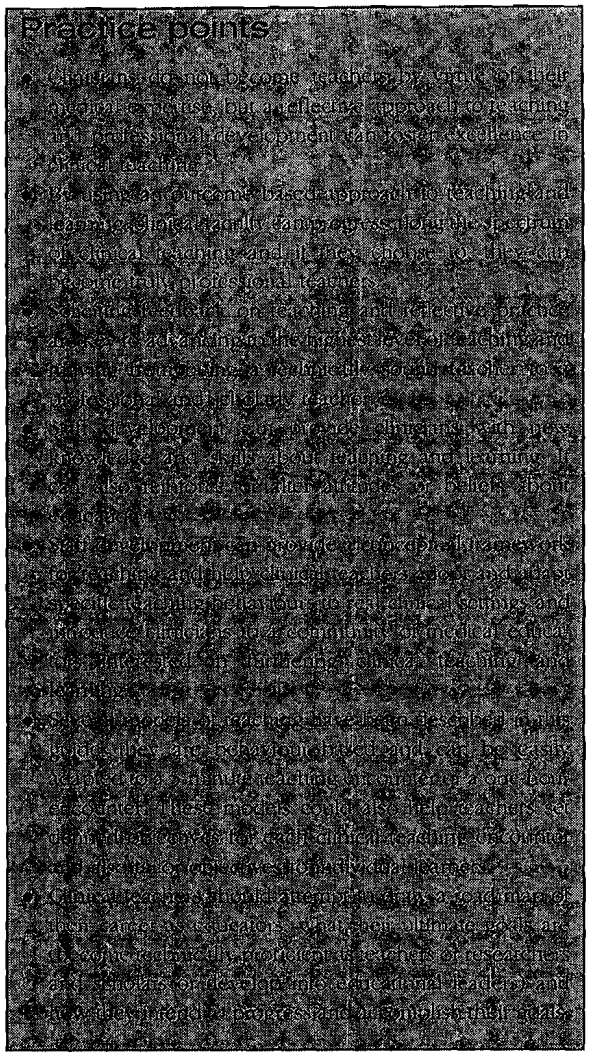
Introduction

Teaching in the clinical environment is defined as teaching and learning focused on, and usually directly involving, patients and their problems (Spencer 2003). The clinical environment consists of inpatient, hospital outpatient and community settings, each with their own distinct challenges. It is in this environment that students learn what it means to be a real doctor. Skills such as history taking, physical examination, patient communication and professionalism are best learned in the clinical setting, medical knowledge is directly applied to patient care, trainees begin to be motivated by relevance and self-directed learning takes on a new meaning (Spencer 2003). Teaching in the clinical setting often takes place in the course of routine clinical care where discussion and decision-making take place in real time. Often the teaching will centre on an analysis of actual patient care that the student has undertaken. This is the most common pattern for postgraduate trainees. Undergraduate students benefit from additional sessions specifically planned for teaching. These sessions may take place in the ordinary clinical environment and make use of the patients who are opportunistically available. They may on the other hand be highly structured with particular patients brought up especially for the session.

The word 'doctor' is derived from the Latin docere, which means 'to teach' (Shapiro 2001). Clinical teachers have a dual role in medicine, to provide patient care and to teach (Prideaux et al. 2000; Irby & Bowen 2004). Though all doctors are usually well prepared for their clinical roles, few are trained for their teaching roles (Steinert 2005). Clinical teachers take their role as teachers of future generations of doctors seriously and with enthusiasm. Yet, most lack knowledge of educational principles and teaching strategies thus may be inadequately prepared for this additional professional role (Wilkerson & Irby 1998). It has simply been assumed that professionals who have graduated from medical schools/colleges and undergone postgraduate training can automatically start teaching the day after they graduate. Due to advances in education such as new methods of teaching and learning, a more student-centred teaching, competency based assessment and emphasis on professionalism; educators today are required to have an expanded toolkit of teaching skills and clinical expertise (Harden & Crosby 2000; Searle et al. 2006).

Correspondence: Dr. Subha Ramani, Boston University School of Medicine, Boston, MA, USA. Tel: 1-617-638-7985; fax: 1-617-638-7905; email: sramani@bu.edu

ISSN 0142-159X print/ISSN 1466-187X online/08/040347–18 © 2008 Informa UK Ltd
DOI: 10.1080/01421590802061613



Box 1. Skills that make a clinical teacher excellent

Excellent clinical teachers:

- share a passion for teaching;
- are clear, organized, accessible, supportive and compassionate;
- are able to establish rapport; provide direction and feedback; exhibit integrity and respect for others;
- demonstrate clinical competence;
- utilise planning and orienting strategies;
- possess a broad repertoire of teaching methods and scripts;
- engage in self-evaluation and reflection;
- draw upon multiple forms of knowledge, they target their teaching to the learners' level of knowledge.

Box 2. Challenges of clinical teaching

- Time constraints
- Work demands – teachers maintain other clinical, research or administrative responsibilities while being called upon to teach
- Often unpredictable and difficult to prepare for
- Engaging multiple levels of learners (students, house officers etc)
- Patient related challenges: short hospital stays; patients too sick or unwilling to participate in a teaching encounter
- Lack of incentives and rewards for teaching
- Physical clinical environment not comfortable for teaching

From Focus group discussions of clinical faculty in the Department of Medicine at Boston University School of Medicine

Clinical teaching overview

What makes a clinical teacher excellent?

Many investigators have examined the qualities that learners value in their clinical teachers. Irby & Papadakis (2001) summarized these and list the skills that make a clinical teacher stand out (see Box 1).

Problems with clinical teaching

John Spencer has listed common problems with clinical teaching in his article on learning and teaching in the clinical environment published in the British Medical Journal's ABC of learning and teaching in medicine series (Spencer 2003). The following are examples of such challenges, though by no means a complete list:

- lack of clear objectives and expectations;
- teaching pitched at the wrong level;
- focus on recall of facts rather than problem solving;
- lack of active participation by learners;
- inadequate direct observation of learners and feedback;
- insufficient time for reflection and discussion;
- lack of congruence with the rest of the curriculum.

Challenges for teachers in the clinical environment

Teaching in the clinical environment comes with its own set of unique challenges (Spencer 2003); some key ones are listed in Box 2.

Despite the numerous challenges noted, many clinicians find practical solutions to overcome them and excel in their dual role as clinician and teacher. The remainder of this guide focuses on practical educational strategies that clinicians can use while teaching in the clinical environment from technical skills to a scientific and professional approach to their teaching.

General teaching models for teaching in any clinical setting

Two models of clinical teaching have been successfully used in faculty development of clinical teachers. Both models are behaviour based and can be adapted by clinical teachers to all clinical settings. The first is the Stanford Faculty Development model for clinical teaching and the second is the Microskills of teaching model, also known as the one-minute preceptor.

Stanford faculty development model for clinical teaching

A popular model for teaching improvement has been the seven-category framework of analysis developed by the Stanford Faculty Development Centre. This comprehensive framework is outlined in the article by Skeff (1988). In addition, this seven-category framework has been validated by work at the University of Indiana which resulted in a 26 item questionnaire that can be used to evaluate teaching (Litzelman et al. 1998). Although it provides a categorical framework for evaluation and analysis of teaching, the power of the model is most effectively demonstrated in hands-on seminars in which faculty are enabled to both understand and apply this method of analysis to their teaching. This model described all clinical teaching as fitting into seven key categories, lists key components under each category and further describes specific teaching behaviours under each key component.

The categories are as follows.

(1) Promoting a positive learning climate: The learning climate is defined as the tone or atmosphere of the teaching setting including whether it is stimulating, and whether learners can comfortably identify and address their limitations. It sets the stage for effective teaching and learning.

(2) Control of session: This refers to the manner in which the teaching interaction is focused and paced, as influenced by the teacher's leadership style. It reflects the group dynamics, which affect the efficiency and focus of each teaching interaction.

(3) Communication of goals: This includes establishment as well as explicit expression of teachers' and learners' expectations for the learners. Setting goals provides a structure for the teaching process, guides teachers in planning the teaching and provide a basis for assessment.

(4) Promoting understanding and retention: Understanding is the ability to correctly analyse, synthesize and apply whereas retention is the process of remembering facts or concepts. This category deals with approaches teachers can use to explain content being taught and have learner meaningfully interact with the content, enabling them to understand and retain it.

(5) Evaluation: It is the process by which the teacher assesses the learner's knowledge, skills and attitudes, based on educational goals previously established. It allows the teacher to know where the learner is and helps them plan future teaching as well as assess effectiveness of teaching. Evaluation can be formative to assess ongoing learner's progress towards educational goals or summative for final assessment to judge learner's achievement of goals.

(6) Feedback: Feedback is the process by which the teacher provides learners with information about their performance for potential improvement. It provides an educational loop through which the teacher can guide learners to use the evaluation of their performance to reassess attainment of goals.

(7) Promoting self-directed learning: Teachers achieve this by facilitating learning initiated by learner's needs, goals and interests. It stresses the importance of acquiring skills to equip the learner to continue learning beyond the time of formal education.

The one-minute preceptor

The 'Microskills' of teaching, also called the one minute preceptor because of the short time available for teaching in the clinical environment, provides a simple framework for daily teaching during patient care (Neher et al. 1992). It is most relevant to teaching postgraduate trainees but the steps also apply to the longer encounters that are specifically focused on teaching for undergraduates. These steps can be used to structure effective short clinical teaching encounters that last five minutes or less as well as to address problems that arise. The original microskills model uses a five-step approach.

Step 1. Getting a commitment: The teacher encourages learners to articulate their opinions on the differential diagnosis and management rather than giving their own conclusions and plans. The teacher must create a safe learning environment so that learners feel safe enough to risk a commitment – even if it is wrong.

Step 2. Probing for supporting evidence: The teacher should encourage learners to 'think out loud' and give their rationale for the commitment they have just made to diagnosis, treatment, or other aspects of the patient's problem. Teachers should either validate learners' commitments or reject them gently if flawed.

Step 3. Teaching general rules: Teachers can guide learners to understand how the learning from one patient can be applied to other situations. The learner is primed for new information they can apply to a given patient as well as future patients. If the learner has performed well and the teacher has nothing to add, this microskill can be skipped.

Step 4. Reinforcing what was done well: It is appropriate to use this microskill every time the trainee has handled a patient care situation well. Effective reinforcement should be specific and behaviour based and not vague. Positive feedback also builds the trainee's self-esteem.

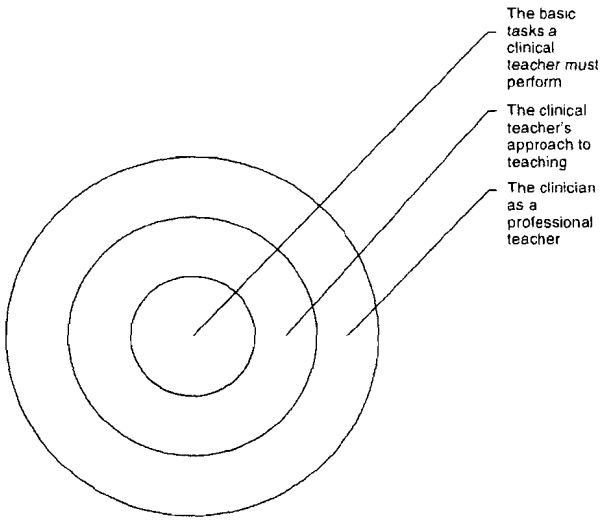
Step 5. Correcting mistakes: Negative or constructive feedback is often avoided by clinical teachers, but this is vital to ensure good patient care. Encouraging self-assessment is a good way to have the learners realise their mistakes themselves and if they have identified their errors, they can be given positive feedback on their self-reflective capabilities. If the teacher has to point out mistakes, this must be specific, timely and based entirely behaviour based.

Applying the Dundee outcomes model in clinical teaching

It has been stated that the medical profession needs to think more seriously about training their teachers and a framework for developing excellence as a clinical educator is needed (Hesketh et al. 2001). Harden et al. (1999) had previously

proposed a 3-circle learning outcomes model to classify skills and abilities that doctors must possess. The Dundee outcomes model offers a user-friendly approach to communicate learning outcomes and was adapted to describe outcomes for medical teachers (Hesketh et al. 2001). We use this model in describing outcomes expected of a clinical teacher, moving from technical competencies to meta-competencies within each circle (Figure 1).

(1) The inner circle refers to the fundamental tasks that clinical teachers should be able to perform competently; doing the right thing.
(2) The middle circle represents the teacher's approach to clinical teaching with understanding and application of relevant learning theories; doing the thing right.
(3) The outer circle represents the development of the individual through a professional approach to teaching in the clinical environment; the right person doing it.



Adapted from Harden et al. 1999

Figure 1. The Dundee 3-circle outcomes model.

In applying the three-circle outcomes model for teachers in the clinical environment we have attempted to keep these outcomes clear and unambiguous, specific, manageable and defined at an appropriate level of generality (Harden et al. 1999) (see Box 3).

Circle one: what the clinical teacher should be able to do (doing the right thing)

We list the following tasks as essential for teachers in the clinical environment: time efficient teaching, inpatient teaching, outpatient teaching, bedside teaching, assessment of learners in the work setting and giving feedback.

Time efficient teaching

Irby & Bowen (2004) described a 3-step approach for time efficient teaching in the clinical environment. All three steps described can be adapted equally well to a one-hour session as a 10-minute teaching session.

Planning. Advanced planning can achieve the following:

- sharpen expectations;
- clarify roles and responsibilities;
- allocate time for instruction and feedback;
- focus learners on important priorities and tasks.

The planning stage includes communicating expectations to learners, soliciting learners' goals, creating a safe and respectful learning environment, selecting appropriate patients for the teaching and priming learners about the goals of the session.

Teaching. Distinguished clinical teachers draw upon a repertoire of teaching strategies to meet the needs of their learners and selectively use any or all of the following five common teaching methods.

- Teaching from clinical cases; combining simple discussions for novice learners with higher level discussions for more senior learners

Box 3. Applying the three-circle outcomes model for teachers in the clinical environment

Tasks of a clinical teacher (Doing the right thing)	Approach to teaching (Doing the thing right)	Teacher as a professional (The right person doing it)
• Time efficient teaching	• Showing enthusiasm for teaching and towards learners	• Soliciting feedback on teaching
• Inpatient teaching	• Understanding learning principles relevant to clinical teaching	• Self-reflection on teaching strengths and weaknesses
• Outpatient teaching	• Using appropriate teaching strategies for different levels of learners	• Seeking professional development in teaching
• Teaching at the bedside	• Knowing and applying principles of effective feedback	• Mentoring and seeking mentoring
• Work based assessment of learners in clinical settings	• Modelling good, professional behaviour including evidence based patient care	• Engaging in educational scholarship
• Providing feedback	• Grasping the unexpected teaching moment	

- Using questions to diagnose not only learners' capacity for recall but also their analysis, synthesis and application capabilities
- Using advanced learners to participate in the teaching
- Using illness and teaching scripts. Examples of illness scripts include knowledge of typical symptoms and physical findings, predisposing factors that place the patient at risk and underlying pathophysiology. Teaching scripts commonly include: key points with illustrations, appreciation of common errors of learners and effective ways of creating frameworks for beginners to build their own 'illness scripts'.
- Acting as role-models at the bedside or in examination rooms

Evaluating and reflecting. Observing learners directly is an important prerequisite for effective feedback. Feedback should be based on observed behaviours, include positive and negative feedback and teachers need to promote self-assessment by learners. These techniques are discussed in greater depth later.

Inpatient teaching

Ende (1997) wrote that the role of the inpatient teacher is one of the most challenging in medical education, that of a master, mentor, supervisor, facilitator, or all of the above. Inpatient teaching can be chaotic and frustrating, as students of varying levels of sophistication and interest fight off (or surrender to) interruptions and urges to sleep, while the attending physician holds forth on unanticipated topics, and about patients who may not be available. Despite the various challenges (see Box 4), he states that inpatient teaching can be riveting if the teachers follow some basic principles. Teachers should try to facilitate knowledge acquisition by asking questions that make learners think and reason rather than recall facts. More importantly, knowledge should be applied to specific patients for clinical problem solving. Teachers should have some knowledge of different learning styles and adapt their teaching style to different learners. Teachers can set a comfortable and safe learning environment in which they and the learners freely ask questions and are prepared to admit their limitations. Inpatient teams also need to behave as a teaching community where each member respects the other in order to maximize their learning. Teachers should learn to challenge their learners without humiliating them and provide support so that learning can be furthered. Ende suggests that in preparation for effective ward teaching, the teachers should ask themselves a set of questions before each teaching encounter.

(1) What do you hope to accomplish?
(2) What is your point of view?
(3) How will your learners be engaged?
(4) How will you meet the needs of each learner?
(5) How will rounds be organized?
(6) Are your rounds successful?
(7) How will you make the time?

Although these questions can be applied to any clinical environment, they are particularly apt for the inpatient setting where a little mental preparation goes a long way. Time constraints, varying learner levels, unexpected teaching moments, presence or absence of the patient can all be factored in while the teacher attempts to answer these questions.

Outpatient teaching

Clinical teaching has recently been moving from the wards to clinics. In recent years, the outpatient clinics have become an integral venue to teach clinical medicine. With shorter hospital stays, it has become impossible for trainees to follow and learn the natural history of a disease from the inpatient environment. Outpatient settings provide one area where trainees can learn this, follow the patient over time and become involved in the psychosocial aspects of patient care (McGee & Irby 1997; Prideaux et al. 2000). Outpatient clinics are exceedingly busy and chaotic settings with very short teacher-trainee interactions (see Box 5). Often, clinical teachers are providing direct

Box 4. Challenges of inpatient teaching

1. Difficult to set teaching goals, unanticipated events occur frequently
2. Ward team usually composed of varying levels of learners
3. Patients too sick or unwilling to participate in the teaching encounter
4. Patient stays are too short to follow natural history of disease
5. Teachers could compromise trainee-patient relationship if they dominate the encounter
6. Trainees and teachers feel insecure about admitting errors in front of the patient and the rest of the medical team
7. Tendency by many clinical teachers to lecture rather than practise interactive teaching
8. Engaging all learners simultaneously can be difficult
9. Teachers need to pay close attention to learner fatigue, boredom and workload

Box 5. Challenges of outpatient teaching

- Busy clinical setting
- Teaching time often short, no time for elaborate teaching
- No control over distribution and organization of time
- Attending to several patients at the same time with multiple learners
- Brief teacher-trainee interactions
- Patient care demands usually take priority and must be addressed
- Multiple patient problems must be addressed simultaneously, so teachers cannot focus on one problem to teach
- Learning and service take place concurrently
- Organic and psychosocial problems are intertwined
- Diagnostic questions often settled by follow up of empiric treatment
- Teacher should be a guide and facilitator than information provider

patient care while supervising and trying to teach students and residents (Neher et al. 1992; McGee & Irby 1997). In a busy clinic, patients too may not be interested in being participants of a trainee-teaching encounter. Overall, service requirements outweigh teaching requirements thus making this an uncontrolled teaching setting. Techniques originally described for effective inpatient teaching do not apply well to outpatient teaching. The outpatient clinic promises many unique educational opportunities including more complete observation of chronic diseases, closer relationships between teachers and learners, and a more appropriate forum for teaching preventive medicine, medical interviewing, and psychosocial aspects of disease (McGee & Irby 1997).

McGee and Irby describe practical tips for efficient teaching in the outpatient settings and they categorize these steps as follows.

(1) Prepare for the visit: Orientate learners of the number of patients to be seen, time to be spent with each patient and how to present patients succinctly.
(2) Teach during the visit: Ask questions to diagnose the learner's knowledge and clinical reasoning, select a specific teaching point in each case, model good physician-patient interactions, observe at least in part learner-patient interactions and provide timely and specific feedback.
(3) Teach after the visit: Answer questions that arise from specific patient problems, clarify what learners did not understand, refer to literature and create reading assignments.

Wolpaw et al. (2003) describe a model for learner-centred outpatient precepting where learners are equal if not the leaders of the teaching interaction. They applied the mnemonic SNAPPS to this model. The six steps of the SNAPPS model are described below.

(1) Summarize briefly history and exam findings: The learner obtains a history, performs an appropriate examination of a patient, and presents a concise summary to the supervisor. The summary should be condensed to relevant information because the preceptor can readily elicit further details if needed.
(2) Narrow the differential diagnosis: For a new patient encounter, the learner may present two or three reasonable diagnostic possibilities. For follow-up or sick visits, the differential may focus on why the patient's disease is active, what therapeutic interventions might be considered, or relevant preventive health strategies.
(3) Analyse the differential diagnosis: In this step, the learner should compare and contrast diagnostic possibilities with evidence of clinical reasoning. This discussion allows the learner to verbalize his or her thinking process and can stimulate an interactive discussion with the preceptor. This discussion also helps clinical teachers to diagnose the level of their learners and thus plan further teaching accordingly.
(4) Probe the preceptor by asking questions about uncertainties, difficulties, or alternative approaches. This step is the most unique aspect of the learner-driven model because the learner initiates an educational discussion by probing the preceptor with questions rather than waiting for the preceptor to initiate the probing of the learner. The learner is taught to utilize the preceptor as a knowledge resource that can readily be accessed.
(5) Plan management for the patient's medical issues. The learner initiates a discussion of patient management with the preceptor and must attempt either a brief management plan or suggest specific interventions. This step asks for a commitment from the learner, but encourages him or her to access the preceptor readily as a rich resource of knowledge and experience.
(6) Select a case-related issue for self-directed learning. The learner may identify a learning issue at the end of the patient presentation or after seeing the patient with the preceptor. The learner should check with the preceptor to focus the reading and frame relevant questions.

Teaching at the bedside

It has been stated that since clinical practice involves the diagnosis and management of problems in patients, teaching of clinical medicine should be carried out on real patients with real problems (Nair et al. 1997). There are many skills that cannot be taught in a classroom, particularly the humanistic aspects of medicine (Nair et al. 1997; Ramani 2003) and require the presence of a patient, real or simulated. The patient's bedside, however, appears to be one of the most challenging settings for clinical teachers. Although many clinical teachers find this an intimidating mode of teaching that bares their own deficiencies, they need to realize that all of them possess a wide range of clinical skills that they can teach their junior and far less experienced trainees (Ramani et al. 2003). Some common sense strategies combined with faculty development programmes at individual institutions can overcome some of this insecurity and promote bedside rounds, which can be educational and fun for teachers and learners alike. Teachers' insecurities can be classified into 2 major domains (Kroenke 2001):

- Clinical domain: Teachers may feel insecure about their knowledge being up to date.
- Teaching domain: Teachers often feel intimidated by having to teach a heterogeneous group of learners who are busy and frequently sleep deprived.

Twelve practical tips have been described to help ease teacher discomfort at the bedside and promote effective bedside teaching (Ramani 2003).

(1) Preparation: Teachers need to familiarise themselves with the clinical curriculum, attempt to diagnose different learner levels and improve their own clinical skills.

(2) Planning: Ende (1997) suggests that all clinical teachers should ask themselves the following questions prior to a teaching encounter and try to answer them:

a. What do you hope to accomplish?
b. What is your point of view?
c. How will your learners be engaged?
d. How will you meet the needs of each learner?
e. How will rounds be organized?
f. Are your rounds successful?
g. How will you make the time?

(3) Orientation: Teachers should obtain objectives of learners, assign roles to each of the team members, try to engage everyone and establish team ground rules.

(4) Introduction: The team of doctors need to be introduced to patients and patients should be oriented about the nature of the bedside encounter; e.g. Patients need to be told that the encounter is primarily intended for teaching and that certain theoretical discussions may not be applicable to their illness.

(5) Interaction: The clinical teachers should serve as role-models during their physician-patient interactions and teach professionalism and a humanistic bedside manner. In addition, teachers should model team work and promote positive team interactions including professional interactions with nursing and other ancillary staff.

(6) Observation: Teachers need not put on a show at the bedside and dominate the bedside encounter (Kroenke 2001). Observing the trainees' interaction with the patient at the bedside can be very illuminating and these observations can be used to plan future teaching rounds.

(7) Instruction: Clinical teachers should avoid asking the trainees impossible questions and 'read my mind' types of questions (LaCombe 1997; Kroenke 2001) and actively discourage one-upmanship among learners. Admitting one's own lack of knowledge might allow trainees to admit their limitations and ask questions. Teachers can role model their willingness to learn by being prepared to learn from trainees.

(8) Summarise: Learners would find it beneficial if teachers summarize what was taught during that encounter. Patients also need a summary of the discussion, what applies and what does not apply to their illness and management.

(9) Debriefing: Time is needed for learners to ask questions and teachers to make clarifications and assign further readings.

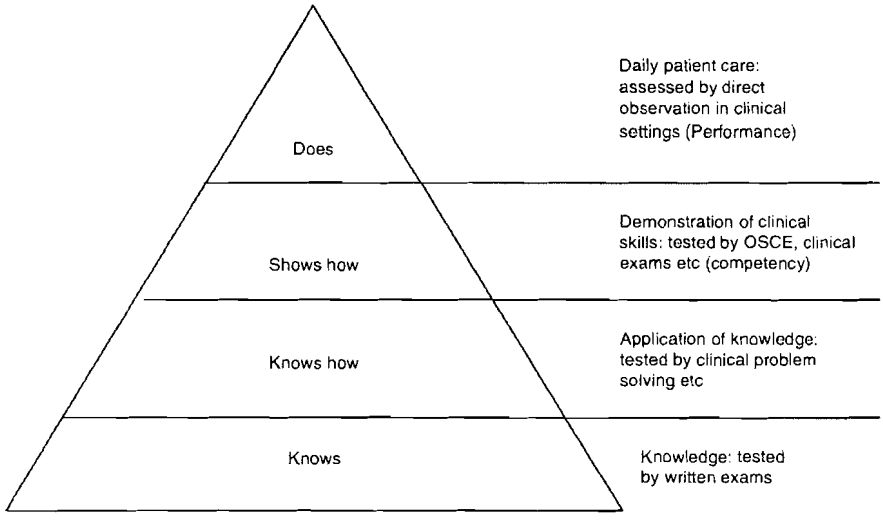
(10) Feedback: Teachers can find out from learners what went well and what did not and give positive and constructive feedback to learners.

(11) Reflection: Reflections about the bedside encounter coupled with learner feedback can help teachers plan the next encounter.

(12) Preparation for the next encounter should begin with insights from the reflection phase.

Work based assessment of learners in the clinical environment

Assessment plays a major role in the process of medical education, in the lives of medical students, and in society by certifying competent physicians who can take care of the public. Society has the right to know that physicians who graduate from medical school and subsequent residency-training programmes are competent and can practise their profession in a compassionate and skilful manner (Shumway & Harden 2003). Miller (1990) proposed his now famous pyramid for assessment of learners' clinical competence (Figure 2). At the lowest level of the pyramid is knowledge



Adapted from Miller (1990)

Figure 2. Miller's pyramid of assessment.

(knows), followed by competence (knows how), performance (shows how), and action (does) The clinical environment is the only venue where the highest level of the pyramid can be regularly assessed.

Studies have indicated that performance in high stakes examinations do not accurately reflect what doctors do in actual patient care (Ram et al. 1999; Rethans et al. 2002). Patient outcomes are the best measures of quality to assess learners in the clinical settings (Norcini 2003), but these are often difficult to ascertain due to factors such as case mix, case complexity, nature of the clinical team and other intangible factors. Assessment in the workplace is quite challenging as patient care takes top priority and teachers have to observe firsthand what the learners do in their interaction with patients and yet be vigilant that patient care is of the highest quality.

Performance outcomes. Norcini (2003) states that the principal measures of performance in the clinical environment include patient outcomes, process of care and volume of services doctors provide.

- Patient care outcomes include morbidity and mortality, physiological outcomes such as blood pressure or diabetes control, clinical events such as stroke or heart attack and last but not least patient satisfaction and experience with care.
- Process of care includes such factors as patient screening, preventive services provided, disease specific measures such as HbA1C for diabetes, aspirin prescription after a heart attack etc.
- Patient volume refers to features such as number of hip replacements performed by orthopaedic surgeons or cardiac catheterizations performed by cardiologists. Volume, in general, correlates with skill and patient morbidity and mortality, but does not always equal high quality patient care.

Clinical teachers should gain familiarity with an outcomes based assessment method appropriate to their own environment (CANMEDS, ACGME, LCME etc.).

Rethans et al. (2002) emphasize that the distinction between competency-based and perfonnance-based methods is important and propose a new model, designated the Cambridge Model, which extends and refines Miller's pyramid. It inverts his pyramid, focuses exclusively on the top two tiers, and identifies performance as a product of competence, the influences of the individual (e.g. health, relationships), and the influences of the system (e.g. facilities, practice time). The model provides a basis for understanding and designing assessments of practice performance.

Assessment methods. In the clinical environment, faculty can readily assess any of the performance measures described above that relate directly to patient care. In these settings, trainees' clinical skills can be assessed outside a simulated or test environment; skills such as patient communication, physical examination, clinical reasoning, case presentation and notes, team work, communication with clinical and non-clinical staff and professionalism. Methods of assessment include examining case records and notes for evidence of diagnostic thinking, listening to case presentations, but the most important method of assessment for clinical teachers would be direct observation. Without observing trainees at work and at the bedside, teachers cannot gather accurate data to provide appropriate feedback.

Giving feedback

In the clinical environment it is vital to provide feedback to trainees as without feedback their strengths cannot be reinforced nor can their errors be corrected (Ende 1983). It is a crucial step in the acquisition of clinical skills, but clinical teachers either omit to give feedback altogether or the quality of their feedback does not enlighten the trainees of their strengths and weaknesses. Omission of feedback can result in adverse consequences, some of which can be long term especially relating to patient care. For effective feedback, teachers need to observe their trainees during their patient interactions and not base their words on hearsay. Feedback can be formal or informal, brief and immediate or long and scheduled, formative during the course of the rotation or summative at the end of a rotation (Branch & Paranjape 2002).

Why is feedback needed? Feedback is essential for a student or intern to gain an insight into what they did well or poorly and the consequences of those actions. If educational goals had been established ahead of the teaching encounter or period, feedback is essential to examine accomplishment or lack thereof of stated goals, re-establish new goals and make action plans to address them (Ende 1983). It tells the learners where they are in comparison to where they ought to be and where they should go. Feedback, when well done, also promotes self-reflection and self-assessment, which are valuable traits for lifelong learning.

Barriers to feedback. One of the biggest hurdles to giving feedback is lack of direct observation of trainees by teachers (Ende 1983). Clinical competence cannot be assessed by written exams, self-report or third party observation, rather this needs to be observed directly by clinical teachers. Teachers are also very hesitant to provide negative feedback and frequently avoid it altogether although this can have adverse consequences on patient care. Trainees, on the other hand, may view negative feedback as a personal attack. Teachers need to establish a positive learning environment in which errors are acknowledged and feedback is expected and accepted. Frequently, feedback is non-specific and unhelpful to learners, e.g. 'good job', 'bad patient communication', etc.

Circle 2: how the clinical teacher approaches their teaching (doing the thing right)

Showing enthusiasm for teaching and towards learners

The starting point for any good teacher must be enthusiasm for the subject being taught. This has to be complemented by an eagerness to transmit this enthusiasm to others, which will necessarily result in a positive attitude to learners.

Enthusiasm for the subject is usually accompanied by a sound knowledge of the subject and a desire to learn more about it, both of which are pre-requisite for successful teaching in higher education. However, while enthusiasm, knowledge and a desire to learn more are necessary for successful teaching they are not sufficient. Teaching is a professional discipline with its own theoretical background and its own recognised techniques. A good teacher must have and apply a working knowledge of both techniques.

Understanding learning principles relevant to clinical teaching

Pedagogy versus androgogy. Much of our approach to teaching and learning is based on studies in children at school and is therefore termed pedagogy. The content of learning is defined by a syllabus and the method of learning is laid down by a curriculum. Both of these may be developed by the individual teacher but are likely to have been laid down by a central authority. The pace of learning is dictated by the teacher. Knowles (1990) studied adults enrolled at evening classes in New York and realised that their approach to learning was different. He coined the term androgogy to cover this approach. The content of the student's studies is dictated by perceived need; the method of learning is selected by the learner and the pace of learning is dictated by the learner. From his observations Knowles derived a set of Principles of Adult Learning which are now widely regarded as crucial to the design of any course for adults (Box 6).

Learning theories. Theories of learning may be neurobiological or behavioural. From a pragmatic educational viewpoint the most useful at present are the behavioural theories. These can be broadly classified as individualistic (based on psychological approaches) and social constructivist (based on sociological approaches). While some of the proponents of each theory will claim that their insights are the only valid approach, the practical educator can draw lessons from all of them. It is important to recognise that the theories are attempting to describe what actually occurs in learning rather than what ought to occur.

Psychological theories. Learning and memory. There is an extensive literature on learning and memory. There appears to be a consensus that different models apply for the learning of knowledge and the acquisition of skills. Clinical teaching must deliver both modalities.

The first stage of acquiring knowledge is the activation of prior knowledge. This is followed by the acquisition of new knowledge. The new knowledge is incorporated into the memory through rehearsal which is more effective if it is done to a third party. The final stage of the learning process is elaboration. This may take the form of transforming the information into a different format e.g. summarizing words as a chart or diagram; comparing and contrasting new information with old, or drawing inferences and conclusions from the total information (Bransford et al. 1999). A slightly different articulation of this process is Schmidt's Information Processing Theory which emphasises the link between the remembering of the new material and the prior knowledge that has been activated which he describes as encoding specificity (Schmidt 1983).

The commonest model used to describe the acquisition of skills is the conscious-competence model. This model is widely used in management training but no-one is entirely clear where it originated. Four stages of ability are described, as described in Box 7.

A fifth stage has been suggested which can be thought of as reflective competence. It is often the case that the person who is operating at the level of unconscious competence is unable to teach others the skill. The person who has reflective competence is able to perform the task without conscious thought but can if necessary analyse what they are doing in order to teach the skill to someone else (Chapman 2007).

Self-determination theory. It is self-evident that students' learning is affected by their motivation. Williams et al. (1999) suggest that the nature of the motivation is important. According to self-determination theory there are two primary kinds of motivation – controlled and autonomous. Controlled motivation is brought about by external pressures (other people's expectations; rewards and punishments) or by internalized beliefs about what is expected. In contrast, autonomous motivation occurs when individuals see the material to be learnt as intrinsically interesting or important. Controlled motivation leads to rote-learning with little

Box 6. Principles of adult learning

Adults:

- have a specific purpose in mind;
- are voluntary participants in learning;
- require meaning and relevance;
- require active involvement in learning;
- need clear goals and objectives;
- need feedback;
- need to be reflective.

Knowles (1990)

Box 7. The conscious-competence model

Unconscious incompetence	The subject is not aware of the skill in question
Conscious incompetence	The subject is aware of the skill and recognizes the need to acquire it
Conscious competence	The subject has acquired the skill but needs to focus their attention on its performance
Unconscious competence	The subject has achieved mastery of the skill and can perform it without conscious thought; other tasks can be performed at the same time.

integration of the material into the student's long term values. Autonomous motivation, among other benefits, leads to greater understanding, better performance, and greater feelings of competence. In addition, students who are encouraged to develop an autonomous approach to learning are more likely to act in ways that promote the autonomy of their patients.

Experiential learning. Most informal learning is based on experience. Kolb (1984) described the process by which this occurs in his learning cycle. Learning occurs when an individual reflects on an experience. On the basis of this reflection, the individual will develop a working theory (although they may not fully articulate it), which will lead them to take a certain course of action. That action will result in a further experience and so the cycle continues with a steady accumulation of useful knowledge. The cycle can be entered at any point. For example, an individual may be told about a theory and take action without having had previous experience of the particular situation. Different individuals will have different preferences for the starting point depending on their learning style (see below).

Sociological theories. Situated learning. Vygotsky, the Russian educational psychologist, postulated on the basis of his study of school children that learning was socially determined and resulted from the interaction of the child with those around her. He observed that if a child has adult or peer support she can solve problems that she is incapable of solving unaided. This difference between aided and unaided performance he called the zone of proximal development and suggested that it is here that learning takes place. In other words, interaction with others is essential to learning (Vygotsky 1978).

Communities of Practice. Clinical activity usually takes place in teams. Such teams are important not only for the delivery of care but for the continuing professional development of the team members. Functional teams form communities of practice in which the individual members support one another. It is a feature of such groups that knowledge and skills are rapidly disseminated throughout the group. This may be through formal structures such as seminars but is more likely to be through the informal day-to-day contact between members. Lave & Wegner (1991) suggest that learners or apprentices are legitimate peripheral participants in such groups. Although they have yet to achieve full membership of the group they are allowed to take part in the activities of the group and in that way they also acquire the knowledge that is inherent in the group. Eventually, they will be absorbed into the group and accepted as a full member of the group. This transition is often marked by ceremony such as passing the final examination.

Reflective practice. At first sight, reflective practice might seem to be an individualistic learning method rather than a social one. However, Schon (1995) identified that reflection is much more effective when conducted with a mentor making it a social activity. He describes two forms of reflection: reflection in action which takes place during an activity, and reflection on action which takes place once that action has been completed. Both are important adjuncts to learning.

Learning styles. It is apparent that different individuals have different approaches to learning. There have been a variety of attempts to describe these different approaches or learning style. Some classifications focus on the cognitive aspects of learning; some focus on the modalities of learning preferred by the learners; a third group focus on the outcomes of the learning.

Cognitive approach – Honey and Mumford Learning Style Inventory. The Honey and Mumford Learning Style Inventory is widely used in management training. It is based on Kolb's learning cycle and identifies four main learning styles (Honey & Mumford 1992).

- Pragmatists prefer to learn directly from experience
- Reflectors prefer to learn by reflecting on their experiences
- Theorists prefer to learn by developing explanations and working theories
- Activists prefer to learn by involvement in activity.

No individual has a single preferred style of learning but each individual will display the learning styles to differing degrees.

Modalities of learning – visual–auditory–kinaesthetic learning style. A potentially more useful learning style questionnaire is the visual–auditory–kinaesthetic (VAK) questionnaire which is widely used in schools. The emphasis is on the subject's preferred modality for acquiring material to be learnt.

- Visual learners prefer material that is delivered through visual media. This includes written and graphic material but also electronic visual media.
- Auditory learners prefer the spoken word to visual material.
- Kinaesthetic learners learn best when the learning involves them in physical activity.

Learners will usually display a mixture of the three learning styles although one may predominate.

Outcomes of learning – deep/superficial learning. The desired goal for learning is that the learners should achieve understanding of the subject. This is called deep learning. When the amount of material to be learnt is too great, or where the assessment of the learning is based purely on recall, learners will display superficial learning. Experienced students will identify those aspects of the material presented which need to be understood for future use and those which will merely need to be recalled for the purposes of assessment. They will adopt a deep learning approach for the former and a superficial learning approach for the latter. This combined approach is described as strategic learning (Newble & Entwhistle 1986).

The teacher's goal must be to develop deep learning. Because students have differing patterns of learning styles, the material to be learnt must be presented in a variety of ways. Patient-centred teaching involves all modalities of the VAK approach as the student will observe the patient, hold conversations with the patient and the instructor and will

carry out physical activity in examining the patient and carrying out clinical procedures. It is also evident that patient-centred teaching will give the student experiences as a result of activities that they undertake. The teacher needs to encourage reflection on what has taken place linked to a discussion of the theoretical background to the case.

Using appropriate learning strategies for different levels of learner

Approaches to teaching in the clinical setting will differ according to the level of the students being taught. Undergraduates are likely to be taught in sessions specifically dedicated to this end. Postgraduate trainees may well be taught in the course of routine service delivery. In any clinical teaching session it is important that the teacher has clear goals and objectives for the session. If the teacher is unsure what they are trying to achieve, the students will not be able to identify the purpose of the session. This will conflict with the principles of adult learning.

Motivation is rarely a problem with students in the clinical setting. Failure to engage with the student is more likely to be a result of poorly constructed teaching sessions rather than student motivation. This is often due to the selection of inappropriate goals for the session.

The purpose of the session will differ depending on the level of the student. The underlying teaching methods can be the same. The new undergraduate who is developing the art of history taking will require different goals from the senior postgraduate student who is learning the nuances of managing variants of the same disease. Both can be taught on the same patient by focusing on different learning tasks. It is not a good idea to try to teach both at the same time as they have different goals and objectives.

The session should start with establishing what the student already knows relevant to the patient's presentation and this should include their understanding of the scientific background as well as the clinical aspects. Failure to establish the students starting point is another common reason for the failure of the student to engage in a teaching session. The topic chosen by the teacher may be too advanced or too elementary for the group of students being taught. In either case the student will have difficulties.

The students should be active participants in the session. Merely telling the students the teacher's view of the situation or having them observe the expert in action does not lead to deep learning. The students should be permitted to carry out relevant components of the clinical task and then be engaged in active discussion. In this way the full range of different learning styles can be accommodated. Dialogue with the student is an important part of clinical teaching. Attention should be paid to probing the students' understanding rather than their simple ability to carry out a mechanical task or recall isolated facts. The questions 'Why' and 'So what' are an essential part of the clinical teachers armamentarium. This will encourage the elaboration stage of learning.

Knowing and applying principles of giving feedback

Feedback should provide the student with the opportunity to reflect on their performance and its possible consequences. It can guide the student's future learning by identifying their strengths and weaknesses (Sender Liberman et al. 2005). The principles of giving feedback have been well-rehearsed by a number of authors. These principles include the use of mutually agreed upon goals as a guide to the feedback; addressing specific behaviours not general performance; reporting on decisions and actions not on one's interpretation of the student's motives; and using language that is non-evaluative and non-judgemental (Ende 1983). These attributes have been found empirically to be valued by trainees (Hewson & Little 1998). Feedback may be corrective (when the student's performance has been inadequate) or reinforcing (well the student has done well) (Branch & Paranjape 2002). Feedback may be formal or informal. In the clinical teaching setting timely, informal feedback is highly valued by the students.

The first requirement of feedback is that the student has a clear concept of the objective they are trying to attain. Feedback can then inform how close they have come to achieving that target and ideally what they need to do differently in order to achieve the target. Direct observation of the performance is necessary if feedback is to be effective. The objective may be a behaviour such as a clinical skill or a cognitive process such as interpreting a history.

At the simplest level feedback informs the student that they have either succeeded or failed at the task. This is common in licensing examinations where the candidate knows either that they have passed or failed but is not told why. In the clinical setting it would be more usual for the student to be told that their performance was inadequate and then a demonstration given of how it should have been done. Once again the student is not offered an analysis of what they did wrong. This approach does not provide the best opportunity for the student to learn and is more akin to evaluation than feedback.

Learning is assisted when both the strengths and the weaknesses of the student's performance are identified and discussed. Feedback is not evaluation and therefore should not use judgemental language or make personal remarks. The emphasis should be on reporting the observed behaviours and thinking and should be detailed and specific rather than general. It is a good technique to start with self-assessment as many astute learners usually identify their errors and the teacher can help make plans to correct those errors and reinforce their strengths. It is often the case that the student's judgement of their performance is harsher than the teacher's and it is important to reassure the student that they have done well.

Clinical learning often takes place in a group environment. In this setting it is helpful to involve the other members of the group in the informal feedback process. They often have valuable insights into why their colleague behaved as she did and, in addition, they will learn the process of constructive feedback. More formal summative feedback should be given in private at a mutually agreed time.

Above all feedback should be constructive. This does not mean that the student's performance cannot be criticized but when there are deficiencies the feedback should include suggestions for making improvements.

Role modelling

An important part of clinical teaching is the development of the professional role in the students. Both trainees (Brownell & Cote 2001) and faculty (Wright & Carrese 2002) agree that the observation of role models is the most important component in this process. This fits well with the theories of situated learning and communities of practise discussed previously. If positive messages are to be transmitted consistently it is essential that teachers reflect on their own attitudes and behaviours (Kenny & Mann 2003). Modelling life long learning requires that the teacher is willing to admit ignorance and prepared to learn from the students. Good doctor-patient relationships and evidence based clinical practice are other areas where the teacher's behaviour will reinforce (or undesirably contradict) their formal teaching.

'Grasping the unexpected teaching moment'

Unpredictability is one of the attractions of clinical practice. There are occasions when it is better to abandon the carefully constructed teaching plan and seize the opportunity which suddenly presents itself. After all, the unexpected will be what excites you and you are likely to transmit that excitement to the students. A sound grasp of the theoretical approaches to teaching are no substitute for enthusiasm for the process of teaching and for the subject that is being taught.

A key prerequisite for using the unexpected teaching moment most efficiently is a teacher's willingness to admit their errors or limitations, thus allowing learners to admit their own without an a climate of humiliation.

Circle 3: the clinician as a professional teacher (the right person doing it)

Even if a teacher can master all the technical competencies listed in the inner circle, emotional and attitudinal competencies such as self-awareness, self-regulation, motivation, empathy and social skills are required to achieve excellence (Harden et al. 1999).

We list the following as essential circle 3 tasks for clinical teachers by which they may become the 'right persons doing it'.

- Soliciting feedback on teaching
- Self-reflection
- Professional development in teaching
- Mentoring.

Soliciting feedback on teaching

Most clinical teachers go about their business of teaching with very little feedback on their strengths and weaknesses as a teacher. Frequently, the only evaluations on their teaching are from learners and these too may be few and far between. Some institutions have adopted a coaching or consulting service for teachers, but these pertain more to classroom teaching or small group teaching rather than teaching in the clinical environment. More institutions should adopt a 360-degree method for evaluating their clinical teachers rather than depend on incomplete and ineffective evaluations from learners alone. These may include measures such as learners' performance and progress as a proxy for teaching impact, video recording of teaching sessions with reflection and feedback, teacher self reports, peer observations etc.

In the face of overwhelming expectations at work, clinical faculty rarely ask for feedback on their teaching from learners or peers. The clinical environment adds an additional twist by the all-important focus on patient care and safety. Thus, frequently the emphasis is on the patient and their management and the teaching strategies are all but forgotten. In the event that a teacher asks their learners for feedback, learners hesitate to offer it as there may be some anxiety about their own evaluations by their teachers. Those learners that offer feedback give non-specific, vague feedback that teachers cannot readily assimilate or apply in their future teaching encounters.

Teachers should be encouraged to seek feedback on their teaching from peers and learners, staff development should train teachers in efficiently obtaining feedback and last but not least a teaching consulting or coaching service developed by institutions for clinical teachers would help improve teaching skills of individual teachers as well as the institution as a whole. Trainees too can benefit from coaching and encouragement on providing useful feedback to their teachers.

Self-reflection

Reflection in medicine has been defined as consideration of the larger context, the meaning, and the implications of an experience or action (Branch & Paranjape 2002) and when used properly allows for the growth of the individual. It has also been stated that professionals must distinguish themselves from technicians by awareness of the larger context of their work using this knowledge for lifelong learning and not limiting themselves to performing specific tasks (Schön 1987; 1983). One might therefore assume that reflection, so essential to educating physicians, is even more crucial for clinical teachers to adopt a professional approach to their teaching, namely be the right person doing it.

Both phases of reflective practice (Kaufman 2003), reflection in action which occurs immediately and reflection on action which occurs later, are readily applicable to clinical teachers.

Fryer-Edwards et al. (2006) have suggested three key teaching skills that illustrate learner-centred, reflective teaching practices and provide a framework for teachers with both cognitive and affective components. Although these teaching practices were developed for communication skills training, they are readily applicable to any clinical environment.

- Identifying a learning edge: Teachers work with learners to identify their learning edge, which is the place where they find learning challenging, but not overwhelming.
- Proposing and testing hypotheses: Teachers formulate hypotheses on issues such as barriers or facilitators to learning for individual learners, learning needs, emotional responses to patients or the rest of the team and apply teaching strategies to test these hypotheses.
- Calibrating learners' self-assessments: This involved learners thinking out aloud their self-assessment, values and beliefs and using these insights to stimulate further reflection.

Professional development

Medical education has traditionally had little input from trained educators. In the past, a high level of clinical competence and experience was considered sufficient to be a good clinician educator, now it is increasingly recognized that teaching itself is a skilled profession. The British General Medical Council in its publication: *Tomorrow's Doctors*, includes the following attributes of a practitioner (General Medical Council 2002).

- Recognition of the obligation to teach others, particularly doctors in training.
- Recognition that teaching skills are not necessarily innate but can be learned.
- Recognition that the example of the teacher is the most powerful influence upon the standards of conduct and practice of trainees.

Most clinical faculty receive little or no explicit training in how to teach, or in theories and processes of teaching. Yet, they are expected to help their trainees master medical knowledge, clinical skills and acquire a habit of lifelong learning. In the changing world of medicine, clinical teachers need to perform time-efficient ambulatory and inpatient teaching, while their own clinical workload keeps increasing. For teachers to succeed at their teaching tasks, faculty development is essential (Wilkerson & Irby 1998). Faculty development also helps teachers build important professional relationships with peers and mentors within and outside their institutions and contribute positively to academic advancement overall (Morzinski & Fisher 2002).

Summary of professional development programmes. Common faculty development formats include train the trainer workshops or seminars, short courses developed by individual institutions, sabbaticals, part time or full time fellowships, scholars programmes and educational workshops at conferences (McLeod et al. 1997; Steinert 1993; 2005; Steinert et al. 2006).

Workshops. The prototypical faculty development programme is a short, focused series of workshops, most of which focus on practical teaching skills development and the educational strategies directly applicable to those teaching skills. Studies demonstrate that such programmes serve a variety of purposes including improving attitudes, self-efficacy, augmenting self-assessed and actual use of specified teaching concepts; facilitating faculty's ability to recognize teaching deficiencies; and increasing knowledge of teaching principles and teaching ability.

Fellowships. In part-time fellowships, faculty spend limited time training at another institution and then work on educational projects at their home institution. Combining the training with the practical application of knowledge and skills at home institutions, such fellowships teach the theory and practice of critical faculty teaching skills. Full-time fellowships are designed to prepare the fellows to be full-time medical educators. Although they include teaching skills, they also emphasize other important educator roles such as educational research and educational leadership.

Teaching scholars programmes. Innovative formats have been developed to link workshops into a more comprehensive programme to target a broader range of outcomes (Gruppen et al. 2003). As a result, some institutions have designed teaching- scholars programmes for their faculty. These programmes are usually a year long and serve as an immersion experience for clinical educators and most require their 'fellow' to complete some educational project. The Teaching Scholars Programme for Educators in the Health Sciences at McGill University (Steinert et al. 2003) and the Medical Education Scholars Programme (MESP) at the University of Michigan Medical School (Gruppen et al. 2003) were designed to create leaders in medical education. These programmes train faculty to provide curriculum design, improved teaching, educational research, and institutional leadership.

Courses at conferences. Many conferences hosted by primary care societies as well as conferences organized by medical education organizations provide a number of courses which focus on teaching and education. These courses range from 90 minute courses to all day courses. Examples of such conferences include the annual conferences of the AAMC, AMEE, Society of General Internal Medicine, Association of Teachers of Family Medicine and the Ottawa conferences.

Co-teaching or peer coaching. In this model, paired physicians focus on developing their teaching skills while sharing the clinical supervision of trainees (Orlander et al. 2000). Through teaching, debriefing and planning, co-teachers gain experience in analysing teaching encounters and develop skills in self-evaluation. Typically, a junior faculty or fellow is paired with a senior faculty educator who helps the 'trainee' teacher reflect on his/her teaching session.

Educational content. The content of staff educational development programmes can be classified under the following key categories.

- Teaching skills: Teaching skills sessions are designed to help participants identify their own needs with respect to teaching skills, and then to practice these skills and receive feedback from colleagues and the faculty (Pololi et al. 2001). Typical topics included in staff educational development include: interactive lecturing, small group discussion, case based teaching, giving effective feedback, promoting a

positive learning climate, communication of goals, evaluation of learners, ambulatory and inpatient teaching and physician patient communication, learner- centred learning, teaching evidence based medicine, stimulating self-directed learning, bedside teaching, etc.

- Educational leadership: This is a higher level of educational development of staff. Having acquired the basic teaching skills, some educators go on to become educational leaders. Examples of topics on the leadership track include: mentoring skills, curriculum development and reform, leadership and management of work teams, running effective meetings, small group leadership, time management, instituting change, cost-effectiveness etc.
- Miscellaneous: Additional skills include learning about instructional technology, using computers in clinical teaching and diversity for the learning environment.

Steinert (2005) has described in depth the reasons and goals of staff development for clinical teachers and also summarized types of professional development resources available.

Mentoring

Several literature reports indicate that mentoring is a useful tool in the academic progression of professionals with many successful academicians attributing their growth and success at least partially to their mentoring relationships. It has also been said that good mentors help their protégés achieve their professional goals more expeditiously. The medical world has well-established research mentoring programmes, but formal mentoring programmes for clinical teachers are scant to non-existent. Mentors can provide guidance, support or expertise to clinicians in a variety of settings and can also help teachers to understand the organisational culture in which they work and introduce them to invaluable professional networks (Walker et al. 2002).

Most successful clinical educators have achieved their success by a trial-and-error approach, seeking multiple senior educators' advice and mentoring on their growth as educators or just talking to their peers. If educating is to be a skilled and scholarly task, educators need mentoring. The ultimate evidence of a clinical teacher being a professional would be if they themselves start mentoring their junior or peer colleagues who wish to achieve professional success in teaching.

Engaging in educational scholarship

For clinical teachers to attain the highest level of professionalism in education and advance academically as educational innovators and leaders, scholarship is essential. Education becomes scholarship when it demonstrates current knowledge of the field, invites peer review, and involves exploration of students' learning. Furthermore, educational work should be made public, available for peer review and reproduced and built on by others (Glassick 2000). Glassick also described six essential criteria of scholarship.

(1) Clear goals
   i. The purpose of the work is clearly stated
   ii. The goals and objectives are realistic and achievable
   iii. The work addresses an important question or need
(2) Adequate preparation
   i. Mastery and understanding of current knowledge in the field and acquisition of skills to carry out the work
   ii. Identifying and obtaining the resources needed to complete the work
(3) Appropriate methods
   i. Using and applying appropriate methods to achieve the stated goals
   ii. Modification of methods to deal with changing circumstances
(4) Significant results
   i. Achievement of the stated goals and objectives
   ii. The work should add to the field and open up additional areas for further exploration
(5) Effective presentation
   i. Using suitable style and organization to present the work at appropriate venues
   ii. Presentation of results with clarity and integrity
(6) Reflective critique
   i. The scholar critically evaluates his or her own work
   ii. The scholar uses evaluations to improve the quality of future work

Points for reflection

(1) How can change be sustained – Change in teaching skills as well as change in attitudes towards teaching?

Other educators have shown that a one-shot approach to educational development does not sustain change and staff development should be longitudinal. Moreover, the educational environment and institutional attitudes towards teaching need to change in order that teaching skills are considered as valuable as research skills in academia.

(2) Can improving clinical teaching skills and excellence in clinical care co-exist? How can teaching initiatives be reconciled with the demands of service?

Clinicians face increasing pressures in their clinical practice and the volume of patients they care for keeps increasing. Time to see patients keeps shrinking and has often been stated as one of the foremost barriers to clinical teaching. Departments and institutions must see high quality teaching as one of their core values; maybe create a core group of faculty who would be responsible for much of the teaching.

(3) Does improvement in clinical teaching matter to patient care? If teaching skills improve, what is the impact on patient management, safety and satisfaction?

This is an area that has not been investigated extensively and is a difficult area to research. Regardless, unless medical educators

Box 8. Practical strategies to achieve Circle-3 clinical teaching outcomes: The teacher as a professional (the right person doing it)

Teaching objectives

Do you establish teaching goals for different types of clinical encounters?

Did you communicate your teaching goals to the learners?

Did you elicit goals of the learners?

Teaching methods

What teaching methods did you use and were they successful (demonstrating, observing, questioning, role-modelling)?

Do you use the same teaching strategies for all learners or do you change your methods for different learner levels and skills?

Feedback

Did you give feedback?

Did you ask for learners' feedback on your teaching?

Planning for the next encounter

Have you used reflective critique of your teaching (from self-assessment or peer or learner feedback) to change your teaching methods?

Professional development

Have you attended courses, studied educational literature or held discussions with other teachers to improve your teaching skills?

Are you planning to engage in the scholarship of teaching, study the impact of your interventions?

demonstrate that improved teaching leads to improved patient outcomes, the public and other stake holders may not see the value of allocating dedicated time to teaching.

(4) How should teachers be evaluated – What outcomes should be measured and who should evaluate them?

Most clinical teachers are evaluated by their trainees, often irregularly and inconsistently. Frequently trainee evaluations are subjective and cursory, thereby of little help to teachers who wish to improve their teaching skills. Trained peers, acting as coaches, may be one of the more useful ways to evaluate teaching, but time needs to be set aside for this coaching model. Microteaching or videotaping of teaching encounters can be invaluable in allowing self-assessment of teaching, but can this be carried out in the clinical environment?

(5) How can institutions and departments elevate the value of clinical teaching – The hidden curriculum, reward its teachers and nurture educational leaders?

In the clinical world, research accomplishments are often held in higher esteem than educational achievements. Expanding academic tracks, staff development, rewarding teachers and establishing clear criteria by which educators can be promoted are possible ways to elevate the value of teaching within institutions and departments.

(6) Teaching clinical skills, bedside teaching – do they really matter? Can technology answer all diagnostic questions?

For better or for worse, technology is here to stay in medicine. Clinical teachers can model appropriate use of technology in making the best clinical decisions and teach trainees the respective value of clinical data and laboratory data in patient care. Educators can further use technology to demonstrate the precision of clinical signs, discarding those that of little value.

(7) How can a clinical teacher set educational objectives when much of the learning is opportunistic? How can teachers respond to the unexpected teaching moments?

Teaching in the clinical environment is beset by frequent unexpected teaching challenges. Questions arise from patients or trainees that teachers are unprepared to answer; patient mood or severity of illness can displace preset teaching objectives. Setting a positive educational environment where teachers are willing to admit their limitations, show willingness to learn from trainees and are prepared to set aside their teaching objectives while grabbing the unexpected moment and doing opportunistic teaching are some strategies to overcome these challenges.

(8) How should teachers inform and orient patients about the teaching nature of the session – Are patients benefiting from the teaching?

If physicians are to learn from direct patient care, patients should be fully engaged in the teaching encounter. Several reports state that most patients enjoy participating in clinical teaching. A few common sense strategies can maximise their impact; introductions, orientation of patients, professionalism, patient education etc, to name a few.

(9) How can a clinical teacher target their teaching to multiple levels of learners and keep them all engaged?

A typical clinical team often consists of multiple levels of trainees from early students to senior house officers and beyond. Clinicians are often intimidated by having to engage all levels during their teaching encounters. Some ways to achieve this successfully include: giving assignments for

trainees to prepare ahead of time, allocating specific tasks at the bedside and using senior trainees to participate in the teaching.

Quotes for Teaching in the Clinical Environment

Summary: Teaching in the clinical environment is a demanding, complex and often frustrating task, a task many clinicians assume without adequate preparation or orientation.

Introduction: Due to advances in education such as new methods of teaching and learning, a more student-centred teaching, competency based assessment and emphasis on professionalism; educators today are required to have an expanded toolkit of teaching skills and clinical expertise

General Teaching models: Two models of clinical teaching have been successfully used in faculty development of clinical teachers. Both models are behaviour based and can be adapted by clinical teachers to all clinical settings.

Stanford Model: Although it (the Stanford model) provides a categorical framework for evaluation and analysis of teaching, the power of the model is most effectively demonstrated in hands-on seminars in which faculty are enabled to both understand and apply this method of analysis to their teaching.

One minute preceptor: The 'Microskills' of teaching, also called the one minute preceptor because of the short time available for teaching in the clinical environment provides a simple framework for daily teaching during patient care.

Applying the Dundee model: It has been stated that the medical profession needs to think more seriously about training their teachers and a framework for developing excellence as a clinical educator is needed.

Time efficient teaching: Irby & Bowen (2004) described a 3-step approach for time efficient teaching in the clinical environment. All three steps described can be adapted equally well to a one-hour session as a 10-minute teaching session.

Inpatient teaching: Ende (1997) wrote that the role of the inpatient teacher is one of the most challenging in medical education, that of a master, mentor, supervisor, facilitator, or all of the above.

Outpatient teaching: In recent years, the outpatient clinics have become an integral venue to teach clinical medicine. With shorter hospital stays, it has become impossible for trainees to follow and learn the natural history of a disease from the inpatient environment.

Teaching at the bedside: It has been stated that since clinical practice involves the diagnosis and management of problems in patients, teaching of clinical medicine should be carried out on real patients with real problems (Nair et al. 1997).

Work based assessment: Assessment plays a major role in the process of medical education, in the lives of medical students, and in society by certifying competent physicians who can take care of the public. Society has the right to know that physicians who graduate from medical school and subsequent residency-training programmes are competent and can practise their profession in a compassionate and skilful manner.

Giving feedback: It (feedback) is a crucial step in the acquisition of clinical skills, but clinical teachers either omit to give feedback altogether or the quality of their feedback does not enlighten the trainees of their strengths and weaknesses.

How the teacher approaches their teaching: The starting point for any good teacher must be enthusiasm for the subject being taught. This has to be complemented by an eagerness to transmit this enthusiasm to others, which will necessarily result in a positive attitude to learners.

Learning and memory: There is an extensive literature on learning and memory. There appears to be a consensus that different models apply for the learning of knowledge and the acquisition of skills. Clinical teaching must deliver both modalities.

Learning styles: It is apparent that different individuals have different approaches to learning. There have been a variety of attempts to describe these different approaches or learning style. Some classifications focus on the cognitive aspects of learning; some focus on the modalities of learning preferred by the learners; a third group focus on the outcomes of the learning.

Knowing and applying feedback: These principles of feedback include the use of mutually agreed upon goals as a guide to the feedback; addressing specific behaviours not general performance; reporting on decisions and actions not on one's interpretation of the student's motives; and using language that is non-evaluative and non-judgemental.

Role modelling: An important part of clinical teaching is the development of the professional role in the students. Both trainees and faculty) agree that the observation of role models is the most important component in this process.

Soliciting feedback on teaching: Teachers should be encouraged to seek feedback on their teaching from peers and learners, staff development should train teachers in efficiently obtaining feedback and last but not least a teaching consulting or coaching service developed by institutions for clinical teachers would help improve teaching skills of individual teachers as well as the institution as a whole

Workshops: The prototypical faculty development programme is a short, focused series of workshops, most of which focus on practical teaching skills development and the educational strategies directly applicable to those teaching skills.

Notes on contributors

DR. SUBHA RAMANI, MD, MMEd, MPH is a general internist and medical educator. She completed Internal Medicine residencies in India and the United States. She has undergone extensive training in medical education including the Stanford Faculty Development Programme, the Dundee Discovery Courses in Medical Education and the Harvard Macy Programme for Physician Educators and completed a Masters in Medical Education from the University of Dundee in 2005. Her major areas of interest are clinical teaching methods, bedside teaching and teaching and assessment of clinical skills. She is the Director of Clinical Skills development for the

Internal Medicine Residency Programme and Associate Professor of Medicine at the Boston University School of Medicine.

PROFESSOR SAM LEINSTER FRCS (Ed), FRCS (Eng) MD is presently the Dean of the School of Medicine, Health Policy and Practice, East Anglia Medical School, and Non-Executive Director and General Surgeon of Norfolk and Norwich Hospital Trust. He was previously Director of Medical Studies and Professor of Surgery at the University of Liverpool where he designed and implemented an acclaimed new curriculum for undergraduate medicine and he has also established the Breast Unit at the Royal Liverpool University Hospital. He is involved internationally in many aspects of medical education and an active member of the Association for the Study of Medical Education (ASME).

References

Branch WT Jr, Paranjape A. 2002. Feedback and reflection: teaching methods for clinical settings. Acad Med 77:1185–1188.

Brandsford J, Brown AL, Cocking RR, editors. 1999. *How People Learn: Brain, Mind, Experience and School* (Washington, National Academy Press).

Brownell AK, Cote L. 2001. Senior residents views on the meaning of professionalism and how they learn about it. Acad Med 76:734–737.

Chapman A. 2007. Conscious competence learning model. Available online at. http://www.businessballs.com/consciouscompetencelearningmodel.htm (accessed 8 October 2007).

Ende J. 1983. Feedback in clinical medical education. Journal of the American Med Assoc 250:777–781.

Ende J. 1997. What if Osler were one of us? Inpatient teaching today. J Gen Intern Med 12(Suppl 2):S41–48.

Fryer-Edwards K, Arnold RM, Baile W, Tulsky JA, Petracca F, Back A. 2006. Reflective teaching practices: an approach to teaching communication skills in a small-group setting. Acad Med 81:638–644.

Glassick CE. 2000. Boyer's Expanded Definitions of Scholarship, the Standards for Assessing Scholarship, and the Elusiveness of the Scholarship of Teaching. Acad Med 75:877–880.

Gruppen LD, Frohna AZ, Anderson RM, Lowe KD. 2003. Faculty development for educational leadership and scholarship. Acad Med 78:137–141.

Harden RM, Crosby JR. 2000. AMEE Guide No 20: The good teacher is more than a lecturer: the twelve roles of the teacher. Med Teach 22:334–347.

Harden RM, Crosby JR, David MH, Friedman M. 1999. AMEE Guide No. 14: Outcome-based education: Part 5-From competency to meta-competency: a model for the specification of learning outcomes. Med Teach 21:546–552.

Hesketh EA, Bagnall G, Buckley EG, Friedman M, Goodall E, Harden RM, Laidlaw JM, Leighton-Beck L, Mckinlay P, Newton R, Oughton R. 2001. A framework for developing excellence as a clinical educator. Med Educ 35:555–564.

Hewson GM, Little ML. 1998. Giving feedback in medical education: verification of recommended techniques. J Gen Intern Med 13:111–116.

Honey P, Mumford A. 1992. *The Manual of Learning Styles* (Maidenhead, Peter Honey Publications).

Irby DM, Bowen JL. 2004. Time-efficient strategies for learning and performance. Clin Teach 1:23–28.

Irby DM, Papadakis M. 2001. Does good clinical teaching really make a difference? Am J Med 110:231–232.

Kaufman DM. 2003. Applying educational theory in practice. Br Med J 326:213–216.

Kenny NP, Mann KV. 2003. Role modelling in physicians' professional formation: reconsidering an essential but untapped educational strategy. Acad Med 78:1203–1210.

Knowles MS. 1990. *The Adult Learner: A neglected species*, 4th edn (Houston, Texas, Gulf Publishing).

Kolb DA. 1984. *Experiential Learning: Experience as a Source of Learning and Development* (Chicago, Prentice Hall, Eaglewood Williams Cliffs).

Kroenke K. 2001. Attending rounds revisited. (President's column). Soc Gen Intern Med Forum 24:8–9.

LaCombe MA. 1997. On bedside teaching. Ann Intern Med 126:217–220.

Lave J, Wenger E. 1991 *Situated Learning: Legitimate Peripheral Participation* (UK, Cambridge University Press).

Litzelman DK, Stratos GA, Marriott DJ, Skeff KM. 1998. Factorial validation of a widely disseminated educational framework for evaluating clinical teachers. Acad Med 73:688–695.

McGee SR, Irby DM. 1997. Teaching in the outpatient clinic. Practical tips. Journal of General Internal Medicine 12(Suppl 2):S34–40.

Mcleod PJ, Steinert Y, Nasmith L, Conochie L. 1997. Faculty development in Canadian medical schools: a 10-year update. Canadian Medical Association Journal 156:1419–1423.

Miller G. 1990. The assessment of clinical skills/competence/performance. Acad Med 65:s63–s67.

Morzinski JA, Fisher JC. 2002. A nationwide study of the influence of faculty development programs on colleague relationships. Acad Med 77:402–406.

Nair BR, Coughlan JL, Hensley MJ. 1997. Student and patient perspectives on bedside teaching. Med Educ 31:341–346.

Neher J, gordon KC, Meyer B, Stevens N. 1992. A five-step 'microskills' model of clinical teaching. J Am Board Family Practitoners 5:419–424.

Newble DI, Entwhistle NJ. 1986. Learning styles and approaches: inplications for medical education. Med Educ 20:162–175.

Norcini JJ. 2003. Work based assessment. Br Med J 326:753–755.

Orlander JD, Gupta M, Fincke BG, Manning ME, Hershman W. 2000. Co-teaching: a faculty development strategy. Med Educ 34:257–265.

Pololi L, Clay MC, Lipkin M JR, Hewson M, Kaplan C, Frankel RM. 2001. Reflections on integrating theories of adult education into a medical school faculty development course. Med Teach 23.276–283.

Prideaux D, Alexander H, Bower A, Dacre J, Haist S, Jolly B, Norcine J, Roberts T, Rothman A, Rowe R, Tallett S. 2000. Clinical teaching: maintaining an educational role for doctors in the new health care environment. Med Educ 34:820–826.

Ram P, Grol R, Rethans JJ, Schouten B, Van der Vleuten C, Kester A. 1999. Assessment of general practitioners by video observation of communicative and medical performance in daily practice: issues of validity, reliability and feasibility. Med Educ 33:447–454.

Ramani S. 2003. Twelve tips to improve bedside teaching. Med Teach 25:112–115.

Ramani S, Orlander JD, Strunin L, Barber TW. 2003. Whither bedside teaching? A focus-group study of clinical teachers. Acad Med 78:384–390.

Rethans JJ, Norcini J, Baron-Maldonado M, Blackmore D, Jolly BC, Laduca T, Lew S, Page GG, Southgate LH. 2002. The relationship between competence and performance: implications for assessing practice performance. Med Educ 36:901–909.

Schmidt HG. 1983. Problem-based learning: rationale and description. Med Educ 17:11–16.

Schon DA. 1987. *Educating the Reflective Practitioner: Toward a New Design for Teaching and Learning in the Professions* (San Francisco, Jossey-Bass).

Schon DA. 1995. *The Reflective Practitioner: How Professionals Think in Action* (New York, Basic Books).

Searle NS, Hatem CJ, Perkowski L, Wilkerson L. 2006. Why Invest in an Educational Fellowship Program? Academic Medicine 81:936–940.

Sender Liberman A, Liberman M, Steinert Y, McLeod P, Meterissian S. 2005. Surgery residents and attending surgeons have different perceptions of feedback. Med Teach 27:470–472.

Shapiro J. 2001. Doctor means teacher. Acad Med 76:711.

Shumway JM, Harden RM. 2003. AMEE Guide No. 25: The assessment of learning outcomes for the competent and reflective physician. Med Teach 25:569–584.

Skeff KM. 1988. Enhancing teaching effectiveness and vitality in the ambulatory setting. J Gen Intern Med 3:S26–S33.

Spencer J. 2003. Learning and teaching in the clinical environment. Br Med J 326:591–594.

Steinert Y. 1993. Faculty development in family medicine. A reassessment. Can Fam Phys 39:1917–1922.

Steinert Y. 2005. Staff development for clinical teachers. Clin Teach 2:104–110.

Steinert Y, Mann S, Centeno A, Dolmans D, Spencer J, Gelula M, Prideaux D. 2006. BEME guide. A systematic review of faculty

development initiatives designed to improve teaching effectiveness in medical education. BEME Guide No. 8. 28:497–526.

Steinert Y, Nasmith L, McLeod PJ, Conochie L. 2003. A teaching scholars program to develop leaders in medical education. Acad Med 78:142–149.

Vygotsky LS. 1978. *Mind and Society: The Development of Higher Mental Processes* (Cambridge, MA, Harvard University Press).

Walker WO, Kelly PC, Hume RF. 2002. Mentoring for the New Millennium. Med Educ. Available online at www.med-ed-online.org (accessed 15 August 2007).

Wilkerson L, Irby DM. 1998. Strategies for improving teaching practices: a comprehensive approach to faculty development. Acad Med 73:387–396.

Williams GC, Saizow RB, Ryan RM. 1999. The importance of self determination theory for medical education. Acad Med 74:992–995.

Wolpaw TM, Wolpaw DR, Papp KK. 2003. SNAPPS: a learner-centred model for outpatient education. Acad Med 78:893–898.

Wright SM, Carrese JA. 2002. Excellence in role modelling: insight and perspectives from the pros. Can J Med 167:638–643.

Journal of Thrombosis and Haemostasis, 6: 1647–1654 DOI: 10.1111/j.1538-7836.2008.03075.x

IN FOCUS

Warfarin for atrial fibrillation in community-based practise

A. J. ROSE,*† A. OZONOFF,‡ L. E. HENAULT† and E. M. HYLEK†

*Center for Health Quality, Outcomes, and Economic Research, Bedford VA Medical Center, Bedford, MA; †Department of Medicine, Research Unit – Section of General Internal Medicine, Boston University School of Medicine, Boston Medical Center, Boston, MA; and ‡Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA

To cite this article: Rose AJ, Ozonoff A, Henault LE, Hylek EM. Warfarin for atrial fibrillation in community-based practise. *J Thromb Haemost* 2008; 6: 1647–54.

See also Fitzmaurice DA. Oral anticoagulation should be managed in the community with treatment aimed at standard therapeutic targets and increased recall intervals. This issue, pp 1645–6.

Summary. *Background:* Previous studies of anticoagulation for atrial fibrillation (AF) have predominantly occurred in academic settings or randomized trials, limiting their generalizability. *Objective:* To describe the management of patients with AF anticoagulated with warfarin in community-based practise. *Methods:* We enrolled 3396 patients from 101 community-based practises in 38 states. Data included demographics, comorbidities, and International Normalized Ratio (INR) values. Outcomes included time in therapeutic INR range (TTR), stroke, and major hemorrhage. *Results:* The mean TTR was 66.5%, but varied widely among patients: 37% had TTR above 75%, while 34% had TTR below 60%. The yearly rates of major hemorrhage and stroke were 1.90 per 100 person-years and 1.00 per 100 person-years. Four percent of patients (n = 127) were intentionally targeted to a lower INR, and spent 42.7% of time with an INR below 2.0, compared to 18.8% for patients with a 2.0–3.0 range ($P < 0.001$). Mean TTR for new warfarin users (57.5%) remained below that of prevalent users through the first six months. Patients with interruptions of warfarin therapy had lower TTR than all others (61.6% vs. 67.2%, $P < 0.001$), which corrected after deleting low peri-procedural INR values (67.0% vs. 67.4%, $P = 0.73$). *Conclusions:* Anticoagulation control varies widely among patients taking warfarin for AF. TTR is affected by new warfarin use, procedural interruptions, and INR target range. In this community-based cohort of predominantly prevalent warfarin users, rates of hemorrhage and stroke were low. The risk versus benefit of a lower INR target range to offset bleeding risk remains uncertain.

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.: +1 617 990 2530; fax: +1 917 591 3104.
E-mail: adamrose@bu.edu

Received 25 March 2008, accepted 26 June 2008

Keywords: anticoagulation, atrial fibrillation, hemorrhage, stroke, warfarin.

Introduction

The number of individuals with atrial fibrillation (AF) in the United States is projected to reach 7.5 million by the year 2020 [1]; many European countries can expect similar increases in the prevalence of AF due to aging of the population. Warfarin has been shown in randomized trials to reduce the risk of stroke in AF by 68% [2]. However, the effectiveness of warfarin is challenged by its variable dose response, narrow therapeutic window, and the need for frequent monitoring of the International Normalized Ratio (INR) [3]. Previous observational studies have explored anticoagulation care predominantly within the setting of large anticoagulation clinics [4–7], but little is known about anticoagulation care in community-based practise.

The objective of the Anticoagulation Consortium to Improve Outcomes Nationally (ACTION) Study was to better define patterns of care in community-based practise related to stroke prevention in AF, including use of lower INR target ranges and their effect on anticoagulation control. Ranges of 1.5–2.5 and 2.0–2.5 have been recommended as a possible strategy to offset bleeding risk, particularly among high-risk elderly patients and high-risk patients receiving antiplatelet agents in addition to warfarin following coronary intervention [8,9]. To date, there are no prospective studies documenting the effect of lower INR targets on percent time in the therapeutic INR range (TTR). However, previous studies do suggest that low INR values may confer an increased risk of stroke without reducing bleeding risk [10,11], raising concerns about this strategy. Other unreported areas of interest included frequency of INR testing, proportion of patients in community practise with stable INR control, time course to achieve control in patients new to warfarin, and effect of interruptions of therapy on INR control. A better understanding of these issues would provide for a more informed interpretation of time-in-range analyses, particularly in an era of quality measurement.

Methods

Patients and practise sites

Methods for the study have been described in detail elsewhere [12,13]. Physician practises that were registered users of CoumaCare® software (Bristol-Myers Squibb, Princeton, NJ, USA) were invited to participate. CoumaCare® was freely available and widely used to assist with patient tracking, data entry, and record keeping, but did not include dosing algorithms or other forms of decision support. CoumaCare® provided a uniformity of data structure among the sites that made our study possible at a time when only 18% of US medical practises had an electronic medical record [14].

In total, 174 practises registered to participate and 101 sites had the technological capability and the review board approval necessary to proceed. All sites had at least one dedicated provider managing warfarin, usually within the setting of a community-based, physician group practise. For all centers, McKesson HBOC BioServices (Rockville, MD, USA), provided on-site training about how to recruit patients, obtain consent, transmit data, and report adverse events in accordance with federal regulatory requirements. McKesson is an independent healthcare services company that provides biomedical support services to the United States Government, industry, universities, and contract research organizations.

Patients were invited to participate by letter, clinic flyer, or in person (at the time of a routine appointment). To be eligible, patients had to be 18 years of age or older and provide written informed consent. In total, 6761 patients were enrolled in the ACTION study; enrollment began in April 2000 and follow-up ended in March 2002. The current report analyzes patients taking warfarin for AF, who represent 50.2% of the entire ACTION cohort.

Data collected included demographic information, indication for anticoagulation, medical diagnoses, INR target range, INR values, warfarin dose, and patient management notes. Missing data fields and data entry errors were flagged and resolved directly with the sites by McKesson HBOC. Any interval of 45 days or more without INR testing or any INR value > 10 or < 0.8 triggered a direct query from the data coordinating center. Resolution of the flag relating to the INR testing interval required validation of continued warfarin use and confirmation that the gap was not related to an adverse event.

The study protocol was approved by the Western Institutional Review Board® (WIRB®) of Olympia, WA, USA, and by local review boards where they existed.

Study variables

Risk factors for stroke were extracted from the anticoagulation electronic record, and a $CHADS_2$ stroke risk score was calculated for each patient [15]. TTR was calculated using linear interpolation, as described by Rosendaal *et al.* [16]. We also used the method described by Fihn *et al.* to calculate INR variability [4,17].

We defined a low INR target range as any range with an upper bound below 2.8. Visit text notes were reviewed for definitive evidence of new warfarin use upon study entry. Patients were considered new to warfarin if they had not yet taken any doses of warfarin as of the first clinic note, or were having their first INR measurement after starting warfarin. Frequency of testing was determined by calculating the mean interval between INR tests for each patient, excluding periods of 120 days or greater between INR measurements (periods when the patient was monitored elsewhere). All intentional interruptions of warfarin therapy for a procedure were documented through review of the anticoagulation clinic notes.

Adverse events

Ischemic stroke, systemic arterial embolism, and major hemorrhage were the adverse outcomes of interest. Major hemorrhage was defined as a fatal event, an event requiring hospitalization with transfusion of at least two units of packed red blood cells, or bleeding involving a critical anatomical site such as the cranium or the retroperitoneum. All patient progress notes were individually reviewed for evidence of adverse events; events were validated directly with the sites by McKesson.

Statistical analyses

We calculated TTR for each patient in the study and described the distribution of TTR among patients. We described the INR control of patients with low INR targets compared to patients with an INR target range of 2.0–3.0. We compared INR control of patients new to warfarin with the INR control of prevalent warfarin users, and described their progression toward stable control by week after inception of therapy. We also investigated the patient-level correlates of better or worse INR control, as measured by TTR.

We divided patients into four groups based on their mean interval between INR measurements and examined the relationship with TTR. We also examined the effect on TTR of excluding INR values within 14 days of a procedure. Finally, we computed unadjusted incidence rate ratios for ischemic stroke/systemic embolus and major hemorrhagic events, stratifying by INR target range, age $\geq$ 80 years, and new warfarin status.

Bivariate comparisons were performed using generalized estimating equations (GEEs), in order to account for intraclass correlation within sites of care. When appropriate, tests for linear trend were conducted across groups. Incidence rate ratios and their 95% confidence intervals (CIs) were conducted using a log–linear model (Poisson regression). Because multiple INR values for each patient violate standard assumptions of independent observations, a comparison of frequencies across INR categories was conducted using Monte Carlo methods. All analyses were performed using the R statistical package, version 2.2 (R Foundation, 2007).

Table 1 Baseline patient characteristics stratified by International Normalized Ratio (INR) target range and prevalent warfarin status

Parameter	Low target INR range (n = 127)	Normal target INR range – new starts (n = 165)	Normal target INR range – prevalent users (n = 3104)
Mean age (SD)	78.1 (8.1)†	72.9 (9.9)	74.0 (9.3)
Female gender	65 (51.2%)	68 (41.2%)	1300 (41.9%)
Race/Ethnicity			
White	120 (94.5%)	153 (92.7%)	2909 (93.7%)
Black	2 (1.6%)	3 (1.8%)	42 (1.4%)
Other	5 (3.9%)	9 (5.5%)	153 (4.9%)
Coronary artery disease	58 (45.7%)*	57 (34.5%)	927 (29.9%)
Stroke risk factors			
Age ≥75	87 (68.5%)*	79 (47.9%)	1662 (53.5%)
Diabetes mellitus	19 (15.0%)	23 (13.9%)	521 (16.8%)
Hypertension	69 (54.3%)	94 (57.0%)*	1455 (46.9%)
Heart failure	27 (21.3%)	32 (19.4%)	722 (23.3%)
Prior stroke	13 (10.2%)	10 (6.1%)	2768 (10.8%)
$CHADS_2$ stroke risk score			
0	12 (9.4%)	27 (16.4%)	483 (15.6%)
1	45 (35.4%)	59 (35.8%)	1087 (35.0%)
2	41 (32.3%)	54 (32.7%)	914 (29.4%)
3	17 (13.4%)	20 (12.1%)	424 (13.7%)
4 or greater	12 (9.4%)	5 (3.0%)	196 (6.3%)

*P < 0.05, compared to prevalent warfarin users. †P < 0.001, compared to prevalent warfarin users.

Results

Demographics and study groups

Of the 6761 total patients enrolled, 3396 were taking warfarin for AF (50.2%) and constituted our study cohort. Patients were drawn from 101 clinical practise sites within 38 states of the United States. Ninety-eight of the sites were community-based office practises (98.5% of patients) and three sites were designated anticoagulation clinics (1.5% of patients). Forty-three percent of the community-based sites were cardiology group practises. There were 2892 total person-years of follow-up. The mean age of participating patients was 74 years and 41.9% were female (Table 1).

There were 127 patients (3.7%) with a low target INR range. These patients were not exclusively targeted below an INR of 2.0; 111 (87%) had an upper bound of 2.5, 2.6, or 2.7 and only 16 (13%) had a target range of 1.5–2.0. Patients with a low INR target range were older (78 vs. 74 years, P < 0.001) and more likely to have coronary artery disease (45.7% vs. 29.9%)

Table 2 Warfarin management and anticoagulation control, stratified by International Normalized Ratio (INR) target range and prevalent warfarin status

Parameter	Low target INR range (n = 127)	Normal target INR range – new starts (n = 165)	Normal target INR range – prevalent users (n = 3104)
Follow-up and INR testing			
Mean days in database [SD]	355 (127)	222 (140)†	332 (131)
Mean number of INR values [SD]	16.2 (8.2)	15.9 (8.3)	16.3 (7.8)
Mean INR values/month [SD]	1.40 (0.56)	2.76 (1.38)†	1.60 (0.83)
Intentional interruptions of warfarin for procedures per 100 patient-months	3.2	4.1	3.5
Frequency of INR by category‡			
1.9 or less	851 (42.8%)	801 (30.9%)	10767 (21.6%)
2.0–3.0	992 (49.9%)	1358 (52.4%)	31059 (62.0%)
3.1–3.9	114 (5.7%)	289 (11.1%)	6399 (12.8%)
4.0 and above	32 (1.6%)	144 (5.6%)	1830 (3.7%)
Time in INR target range			
Time below range (%)	42.7%†	27.9%†	18.8%
Time in range (%)	52.5%†	57.5%†	67.5%
Time above range (%)	4.8%†	14.6%	13.6%
INR variability			
Mean INR value	2.15†	2.41*	2.47
Standard deviation	0.53*	1.00†	0.68

*P < 0.05, compared to prevalent warfarin users. †P < 0.001, compared to prevalent warfarin users. ‡P < 0.001 via Monte Carlo.

compared to patients with an INR target range of 2.0–3.0. Patients newly starting warfarin (n = 165) were similar to prevalent users in terms of age and $CHADS_2$ score.

INR control in patients with a low target INR range

Patients with INR targets below the standard range spent 42.7% of time with an INR below 2.0 (Table 2), compared to 18.8% for patients with the standard target range ($P < 0.001$). The mean INR value among patients with a low target range was 2.15, significantly lower than patients with a standard target range (2.47). Conversely, the low target group spent less time with an INR greater than 3.0 (4.8% vs. 13.6%, $P < 0.001$).

INR control in patients new to warfarin

Compared to prevalent warfarin users with normal target INR ranges, patients new to warfarin (Table 2) had lower TTR (57.5% vs. 67.5%, $P < 0.001$), more time in the subtherapeutic range (27.9% vs. 18.8%, $P < 0.001$), and underwent more frequent testing (2.76 INR measurements per month vs. 1.60, $P < 0.001$). Table 3 shows the time to stable INR control over the first six months of therapy, compared with prevalent warfarin users. At six months, the INR control of the new warfarin group had yet to match that of the prevalent warfarin group.

Time in therapeutic INR range among the entire cohort

As is customary when calculating TTR [16], we did not interpolate between INR values separated by more than 56 days; 5.2% of person-time was not interpolated for this reason. For the entire cohort, the mean TTR was 66.5% (standard deviation 19.9%; median 68.3%, interquartile range 54.7–81.2%). TTR varied greatly among patients: 37% of the cohort achieved 75% or greater TTR ('excellent'), 29% had a TTR between 60% and 75% ('good'), and 34% had a TTR of 60% or less ('poor').

Table 3 Percentage of International Normalized Ratio (INR) values in the target range by consecutive week and month of warfarin therapy. The first four weeks are calculated separately only for the new starts group; beyond that, results are tabulated using 28-day months

	New starts (n = 165)		All others (n = 3104)	
Week of study	Number of INR values	% INR values in-range	Number of INR values	% INR values in-range
Index INR	165	21.2	3104	60.4
Week 1	200	40.0	–	–
Week 2	166	48.2	–	–
Week 3	127	51.2	–	–
Week 4	111	55.0	–	–
Month 2	324	56.5	4106	61.3
Month 3	258	58.1	3888	61.1
Month 4	224	54.0	3852	62.9
Month 5	175	55.4	3618	63.3
Month 6	146	54.1	3498	62.2

Table 4 Descriptive statistics for patients with poor (< 60%), good (60–75%), and excellent (> 75%) results for time in the therapeutic International Normalized Ratio range (TTR)

Parameter	TTR < 60% (n = 1141)	TTR 60%–75% (n = 1009)	TTR > 75% (n = 1246)
Mean age (SD)	73.8 (9.9)	74.5 (9.1)	74.2 (8.8)
Female gender	46.5%†	42.6%*	37.9%
Race/Ethnicity			
White	91.9%*	93.6%	95.4%
Black	2.4%*	1.0%	0.8%
Other	5.7%	5.5%*	4.7%
Coronary artery disease	32.4%*	31.6%	28.3%
Stroke risk factors			
Age ≥75	53.5%	54.7%	53.4%
Diabetes mellitus	17.5%	16.4%	15.9%
Hypertension	47.2%	48.0%	47.8%
Congestive heart failure	26.1%†	22.5%	20.5%
Prior stroke	12.0%	9.1%	10.4%
$CHADS_2$ stroke risk score			
0	14.9%	15.8%	15.5%
1	33.0%	34.7%	37.2%
2	30.6%	29.7%	28.9%
3	13.9%	14.9%	12.2%
4 or greater	7.5%	5.0%	6.2%

*$P < 0.05$, compared to > 75% group. †$P < 0.001$, compared to > 75% group.

We compared demographics and comorbidities among these three groups with excellent, good, and poor TTR (Table 4). Mean age and $CHADS_2$ risk score for stroke were similar in the three groups, but more patients with heart failure and/or coronary artery disease were among those with poor control. The group with excellent INR control had the lowest proportion of females (37.9%), followed by the good control group (42.6%), with the highest proportion of females in the poor control group (46.5%; $P < 0.001$ for trend).

Frequency of INR testing

The mean interval between INR tests among all patients was 22.2 days (standard deviation 7.4). The relationship between

Table 5 Time in therapeutic International Normalized Ratio (INR) range (TTR) and INR variability compared among groups stratified by mean interval between INR tests

Mean interval between INR tests (days)	Number of patients (%)	TTR (95% CI)*	INR variability (95% CI) *
< 14	446 (13)	51.6% (49.7, 53.5)	1.22 (1.10, 1.33)
14–20.99	1041 (31)	62.8% (61.7, 63.8)	0.90 (0.85, 0.96)
21–27.99	1197 (35)	69.9% (68.9, 70.9)	0.53 (0.50, 0.55)
> 28	712 (21)	75.6% (74.1, 77.1)	0.33 (0.31, 0.35)

*$P < 0.001$ for trend for both TTR and INR variability.

testing interval and INR control is illustrated in Table 5. A longer interval between tests was strongly associated with increased TTR and reduced INR variability ($P < 0.001$ for trend for both). In particular, patients with the longest intervals had the best control: 21% of patients averaged greater than 28 days between INR measurements and had a mean TTR of 75.6%.

TTR and intentional interruptions of warfarin

Twenty-eight percent ($n = 946$) of patients had at least one intentional interruption of warfarin for a procedure. A subset of these patients ($n = 425$, 13%) had at least one INR below 1.5 recorded within 14 days of an interruption. These 425 patients had a mean TTR of 61.6% compared to the mean TTR of 67.2% for the remainder of the study cohort ($P < 0.001$). After deletion of all INR values within 14 days of a procedure, mean TTR in the two groups was 67.0% and 67.3%, no longer a statistically significant difference ($P = 0.32$).

Rates of adverse events

There were 55 major hemorrhagic events during 2892.1 person-years of observation, providing an incidence rate of 1.90 events/100 person-years (95% CI 1.46–2.48). Six (11%) events were intracranial hemorrhages and 36 (65%) were gastrointestinal hemorrhages. Major hemorrhage (Table 6) occurred more frequently among patients who were ≥80 years of age [unadjusted incidence rate ratio (IRR) 2.07; 95% CI 1.24–3.44] and among those with TTR less than 60% (unadjusted IRR 2.37; 95% CI 1.17–4.79). The low INR target group experienced nearly twice the rate of major hemorrhage despite spending significantly less time with an INR greater than 3.0; however, this difference was not statistically significant. Patients new to warfarin also appeared to be at higher risk of bleeding, but the difference was not statistically significant.

There were 29 ischemic stroke/systemic embolic events among 2892.1 person-years, an incidence rate of 1.00 event/100 person-years (95% CI 0.70–1.44). A limited number of events precluded a calculation of risk in subgroups. Of interest, among patients with a low target INR range, there were two stroke/embolic events in 111.6 patient-years, giving a crude incidence rate of 1.79 events/100 person-years. The unadjusted incidence rate ratio for such patients, compared to all others, was 1.85, but the difference was not statistically significant (95% CI 0.42–6.60).

Table 6 Incidence rates (IRs) and incidence rate ratios (IRRs) for major hemorrhage

Risk factor	Patient-years of follow-up	Major bleeds	IR per 100 person-years	IRR (95% CI) unadjusted
Patient new to warfarin				
No	2798.2	53	1.89	–
Yes	93.9	2	2.13	1.11 (0.27, 4.63)
Low target range				
No	2780.6	51	1.83	–
Yes	111.6	4	3.58	1.88 (0.66, 5.32)
Age ≥ 80				
No	2015.0	29	1.44	–
Yes	876.4	26	2.97	2.07* (1.24, 3.44)
TTR ranges				
TTR > 75%	1093.4	14	1.28	–
TTR 60–75%	917.3	14	1.53	1.20 (0.51, 2.81)
TTR < 60%	881.4	27	3.06	2.37* (1.17, 4.79)

*$P < 0.05$ for comparison. TTR, time in therapeutic International Normalized Ratio range.

Discussion

Summary of main results and TTR variation among subjects

We report the results of a large, nationally representative, community-based cohort study of anticoagulation management and outcomes in patients with AF. The mean TTR in our sample (66.5%) is similar to that achieved in clinical trials in a recent meta-analysis (66.4%) [18]. TTR varied widely among individual patients in our study, with approximately one-third of patients below 60% and one-third above 75%. This distribution is strikingly similar to that recently reported from the warfarin-treated patients in the SPORTIF (Stroke Prevention Using an Oral Thrombin Inhibitor in Atrial Fibrillation) trials: 33% below 60%, 33% 60–75%, and 33% above 75%. These three groups were found to have different rates of adverse events such as hemorrhage and stroke [19]. Similarly, in our study, the group with the worst control had a higher rate of major hemorrhage than the group with the best control (unadjusted IRR 2.37).

The wide variability of TTR among patients in our study suggests that some patients have an easier time achieving and maintaining stable anticoagulation than others. It is important to note that age was essentially the same across these TTR groups, suggesting that the increased rate of hemorrhage among elderly patients is not a result of more erratic INR control. The finding that male gender is associated with higher TTR remains to be explored. Further study is needed to elucidate the ability of high-quality clinical management to minimize variability of TTR among patients.

Patients with low target INR range

One of our most striking results is that patients with a low target INR range spent 42.7% of time with an INR below 2.0. As noted earlier, only 13% of these patients had a designated target of 1.5–2.0. Despite less time with an INR > 3.0, these patients seemed to have a higher rate of bleeding than patients with normal target ranges, although this difference did not reach statistical significance. However, this reinforces the fact that these patients were indeed at elevated risk for bleeding, and underscores the complexity of managing anticoagulation in such patients. There also seemed to be a higher rate of

thromboembolic events in the low INR target group, but the difference did not achieve statistical significance, probably because of a limited number of events. However, we did show that patients with low target INR ranges spend a great proportion of time with an INR below 2.0. Given the known increase in the risk of stroke when the INR is below 2.0 [10,11], patients and clinicians need to be cognizant of the potential trade-offs inherent to the use of a lower INR target range for stroke prevention in AF.

Patients new to warfarin

Our study helps to characterize the natural history of INR control in patients who are new to warfarin and suggests a persistent difference compared to prevalent warfarin users, even at six months post-initiation. This finding emphasizes the adherence and survivor biases that are intrinsic to longer-term use of warfarin and thus to cohorts of prevalent users.

Frequency of INR monitoring

We found that longer INR monitoring intervals were associated with improved INR control. In the United States, clinical guidelines currently recommend that INR testing occur at least every 28 days for all patients [3,9], but such recommendations are based on expert consensus rather than evidence. Our results suggest that clinicians are able to identify patients who can safely be tested less often, and call into question whether all patients must be tested with a fixed minimum frequency in clinical practise. Indeed, despite testing INR as seldom as every 12 weeks [20], British patients achieve TTR results of 65–70% in usual practise [21,22]; these results are at least as good as those achieved with the more frequent testing intervals common in US practise. Rather than a fixed maximum recall interval, recall intervals might be tailored to recent INR control [17,23]. Some software programs already provide optimized recall intervals based on recent INR control, most notably the Birmingham Anticoagulation Program for Primary Care (BAP-PC) [24].

Effect of interruptions of warfarin therapy on measurement of TTR

We found that measurement of TTR can be affected by low INR values because of intentional interruptions of therapy. In our sample, the deletion of INR values for 14 days before and after a procedure increased TTR from 62% to 67%, a clinically meaningful difference. This refinement improves the validity of TTR as a measurement of the quality of anticoagulation care, because the documentation of low INR values proximal to procedures does not imply poor care.

Rates of stroke and hemorrhage

The rates of adverse events in our study, 1.90 major hemorrhage and 1.00 stroke/systemic embolus per 100 person-years, are similar to other, relatively recent studies of patients with AF taking warfarin. For example, among a large cohort of patients taking warfarin for AF from Kaiser Permanente Northern California, Go *et al.* [7] found rates of 1.52 and 1.17 for major hemorrhage and stroke/systemic embolus, respectively. However, it should be noted that any cohort of predominantly prevalent warfarin users, such as ours, is enriched with patients able to tolerate warfarin therapy without complications, while patients new to warfarin may experience higher rates of adverse events [25].

Limitations

Our study has several limitations. First, the small number of stroke/systemic embolic events precluded a calculation of risk in key subgroups of patients. However, previous studies have already quantified the risk of stroke among many important subsets of patients [15]. Secondly, it is possible that healthier patients preferentially gave informed consent to participate in this study. However, the distribution of stroke risk factors and rates of adverse events (stroke, major hemorrhage) are similar to those found in previous studies [7], suggesting that our patients are representative. In addition, our finding that 34% of patients spent 60% or less time in the therapeutic range argues against biased patient selection.

Conclusions

Anticoagulation control in our cohort of community-based patients with AF was similar to participants in clinical trials, but the TTR of individual patients varied widely. Further studies are needed to investigate patient-level and site-level determinants of TTR. Our results concord with those of previous studies [7], suggesting that among prevalent users of warfarin, rates of major hemorrhage and ischemic stroke are low. Our study also highlights the complexity of antithrombotic management among patients deemed to be at highest risk of hemorrhage. Optimal management of these patients has yet to be determined.

Acknowledgements

This study was funded by Bristol-Myers Squibb, the makers of Coumadin® brand warfarin. Bristol-Myers Squibb had no role in the design and conduct of the study, or in the collection, analysis, and interpretation of the data, or in the preparation, review, and approval of the manuscript. Dr Rose is supported by a career development grant from the United States Department of Veterans Affairs, Health Services Research and Development Service. The opinions expressed in this manuscript do not necessarily represent the views or policies of the Department of Veterans Affairs.

Disclosure of Conflict of Interests

The authors state that they have no conflict of interest.

Appendix

The following practises and directors participated in the study (the sites are listed in decreasing order of number of patients enrolled): Lutheran General Hospital, Niles, Illinois—W. Fried, M. Pubentz; Physicians, Inc., Lima, Ohio—D. Parker; Idaho Cardiology Associates, Boise, Idaho—F. Badke; North Clinic, Robbinsdale, Minnesota—V. Krug; Rockwood Clinic P.S., Main, Spokane, Washington—J. S. Pennock; Wenatchee Valley Clinic, Wenatchee, Washington—R. Kirby Primm, L. Vaughn; Framingham Heart Center, Framingham, Massachusetts—J. Dangel, S. R. Hewett; Clinic Pharmacy Consultants-Brainerd Medical Center, Brainerd, Minnesota—B. Twamley, R. Sorenson; Woodland Healthcare, Woodland, California—L. Smith, T. Fajerson; Cardiology, PC, Syracuse, New York—S. O'Donnell; Health Care American Corp, Bradenton, Florida—C. Hoffman; DuPage Medical Group, Department of Cardiology, Winfield, Illinois—N. Kinsley; Camino Medical Group, Sunnyvale, California—S. Edwards; Ohio Valley Heartcare, Evansville, Indiana—L. Janeira, J. Robb; Desert Medical Group/Oasis IPA, Palm Springs, California—H. F. Bellaci, J. Bellaci; Anchor Health Center, Naples, Florida—M. Means; Sutter Gould Medical Foundation, Modesto, California—J. E. Baker; Hannibal Clinic Inc., Hannibal, Missouri—L. Chalton; Saratoga Cardiology, Saratoga Springs, New York—R. Sheldon, D. Kandath; Lima Memorial Hospital, Lima, Ohio—C. L. Thompson, J. Recker; Staten Island University Hospital, Staten Island, New York—M. Howard; Jacksonville Cardiovascular Clinic, Jacksonville, Florida—R. A. Benson; River Valley Healthcare, Silvis, Illinois—K. Carroll; Family Physician Incorporated, N. Canton, Ohio—H. Marshall; Internal Medicine of Northern Michigan, Petoskey, Michigan—P. D. Blanchard; Redmond Internal Medicine, Redmond Oregon—D. Palmer, C. Gangan; Grove Hill Medical Center, New Britain, Connecticut—M. S. Werner; Olean Medical Group, Olean, New York—H. D. Storch, T. L. Buzzard; Internal Medicine Associates of Greenville, Greenville, South Carolina—J. S. Moore; Magan Medical Clinic, Covina, California—R. Sakamoto; Owatonna Clinic—Mayo Health System, Owatonna, Minnesota—T. Price; Dearborn Cardiology, Dearborn, Michigan—S. Dabbous; Westchester Medical Group, White Plains, New York—B. Newman; Central Cardiology Medical Clinic, Bakersfield, California—W. Nyitray; Salem Clinic, Salem, Oregon—M. Smith; East Carolina University, Greenville, North Carolina—C. Estrada; Northwest Primary Care Group, Milwaukie, Oregon—D. McAnulty, P. Devisser; The William W Backus Hospital, Norwich, Connecticut—S. Johnson; Jefferson City Medical Group, Jefferson City, Missouri—C. Balcer; Saint Louis University Department of Neurology, St Louis, Missouri—S. Cruz-Flores, E. Holzemer; Wellspan Health, Yorktowne, York, Pennsylvania—J. D. Horton; Mercy Medical Center, Canton, Ohio—M. Cudnik; Cardiovascular Group, Lawrenceville, Georgia—B. Craig-Allen; Asheville Cardiology Assoc, Asheville, North Carolina—W. Wharton, A. Moser; Cardiac Consultants Chartered, Bethesda, Maryland—L. Chappell; Valley Care Health System, Pleasanton, California—N. Huynh; Bloomington Hospital, Bloomington, Indiana—K. Kalotta; Samaritan Anticoagulation Service, Corvallis, Oregon—R. Stockberger; Covenant Clinic, Waterloo, Iowa—D. Kohls; Dartmouth-Hitchcock Nashua, Nashua, New Hampshire—L. Cook; Cardiology Consultants, PC, Hamden, Connecticut—A. M. Radoff; Seventh Avenue Family Health Center, Fort Lauderdale, Florida—J. Berges; Diagnostic Cardiology, P.A., Jacksonville, Florida—P. D. Kuhlman; Norlanco Medical Associates, Elizabethtown, Pennsylvania—J. Rittenhouse; University of Texas Medical Branch, Galveston, Texas—H. von Marensdorff; Bend Memorial Clinic, Bend, Oregon—M. Hegewald; Memorial Primary Care Center, Hollywood, Florida—J. Beck; Batey Cardiovascular Center, Bradenton, Florida—D. Calabrila, E. J. Sanchez; Western Montana Clinic, Missoula, Montana—W. B. Bekemeyer, D. Ramsey; Winona Clinic, Winona, Minnesota—L. Tschumper; Cardiac Consultants, Lancaster, Pennsylvania—M. Lesko; Hattiesburg Clinic, Hattiesburg, Mississippi—A. J. Jackson; Bryn Mawr Medical Specialist Association, Bryn Mawr, Pennsylvania—H. Mayer; River Valley Healthcare, Moline, Illinois—B. Cady; Cardiovascular Group, Snellville, Georgia—L. Lesser; Medicor, Bridgewater, New Jersey—P. Saulino, C. Hartpence; Bond Clinic, P.A., Winter Haven, Florida—P. Lundsford. K. Bhatia; University of Cincinnati-Pharmacy Anticoagulation Services, Cincinnati, Ohio—J. McQueen; Senior Healthcare Center, Gainesville, Florida—M. L. Breeser; North Canton Medical Foundation, North Canton, Ohio—H. M. Schenker; Manor Family Health Center, Millersville, Pennsylvania—J. Ichter; Cardiology Associates of Central Florida, Ocala, Florida—L. McDaniel; Cardiovascular Associates Ltd, Chesapeake, Virgina—S. R. Jones; Woodburn Medical Clinic, Woodburn, Oregon—F. Golden; Rockwood Clinic North, Spokane, Washington—C. Laudenbach, J. S. Pennock; Wachspress, Shatkin & Rainear, Vineland, New Jersey—L. Assink; Chambersburg Hospital, Chambersburg, Pennsylvania—D. Grant; Wellspan Pharmacy-Dallastown, Dallastown, Pennsylvania—T. G. Williams; Pulmonary & Critical Care Associates, Ypsilanti, Michigan—W. F. Patton; Island Cardiac Specialist, Mineola, New York—P. Ragno; Portland Cardiovascular Institute 2, Portland, Oregon—R. Chelfky; River Valley Healthcare ACS, Bettendorf, Iowa—W. Langley; Consultants in Cardiology, Farmington Hills, Michigan—G. M. McKendrick; Portland Cardiovascular Institute 1, Portland, Oregon—R. Chelfky; Cleveland Clinic Florida, Weston, Florida—B. Fernandez; BiState Medical Consultants, St Louis, Missouri—P. M. Stein, C. B. Lomnel; Medical Consultants, PC, Muncie, Indiana—J. Bow; Cardiovascular Associates of South Florida, Coral Gables, Florida—J. S. Palmer; Parkway Cardiology Associates, Oak Ridge, Tennessee—S. Cooke; Northwest Georgia Diagnostic Clinic, Gainesville, Georgia—J. Jackson; Cardiovascular Associates, Kingsport, Tennessee—L. H. Cox; Heart Place, Dallas, Texas—C. N. Bowers; Rockwood Clinic, Spokane, Washington—C. Laudenbach; J. S. Pennock; Delaware Heart Group, Newark, Delaware—C. Bowens; Abilene Diagnostic Clinic, Abilene, Texas—P. Howard.

References

1 Miyasaka Y, Barnes ME, Gersh BJ, Cha SS, Bailey KR, Abhayaratna WP, Seward JB, Tsang TS. 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**: 119–25.

2 Hart RG, Pearce LA, Aguillar MI. Meta-analysis: antithrombotic therapy to prevent stroke in patients who have non-valvular atrial fibrillation. *Ann Intern Med* 2007; **146**: 857–67.

3 Ansell J, Hirsh J, Poller L, Bussey H, Jacobson A, Hylek E. The pharmacology and management of the vitamin K antagonists: the Seventh ACCP Conference on Antithrombotic and Thrombolytic Therapy. *Chest* 2004;**126**(3 Suppl.):204S–33S.

4 Fihn SD, McDonell M, Martin D, Henikoff J, Vermes D, Kent D, White RH. Risk factors for complications of chronic anticoagulation. A multicenter study. Warfarin Optimized Outpatient Follow-up Study Group. *Ann Intern Med* 1993; **118**: 511–20.

5 Palareti G, Leali N, Coccheri S, Poggi M, Manotti C, D'Angelo A, Pengo V, Erba N, Moia M, Ciavarella N, Devoto G, Berrettini M, Musolesi S. Bleeding complications of oral anticoagulant treatment: an inception-cohort, prospective collaborative study (ISCOAT). Italian Study on Complications of Oral Anticoagulant Therapy. *Lancet* 1996; **348**: 423–8.

6 van der Meer FJ, Rosendaal FR, Vandenbroucke JP, Briet E. Bleeding complications in oral anticoagulant therapy. An analysis of risk factors. *Arch Intern Med* 1993; **153**: 1557–62.

7 Go AS, Hylek EM, Chang Y, Phillips KA, Henault LE, Capra AM, Jensvold NG, Selby JV, Singer DE. Anticoagulation therapy for stroke prevention in atrial fibrillation: how well do randomized trials translate into clinical practice? *Jama* 2003; **290**: 2685–92.

8 Anderson JL, Adams CD, Antman EM, Bridges CR, Califf RM, Casey DE Jr, Chavey WE II, Fesmire FM, Hochman JS, Levin TN, Lincoff AM, Peterson ED, Theroux P, Wenger NK, Wright RS, Smith SC Jr, Jacobs AK, Halperin JL, Hunt SA, Krumholz HM, *et al.* ACC/AHA 2007 guidelines for the management of patients with unstable angina/non ST-elevation myocardial infarction: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation* 2007; **116**: e148–304.

9 Fuster V, Ryden LE, Cannom DS, Crijns HJ, Curtis AB, Ellenbogen KA, Halperin JL, Le Heuzey JY, Kay GN, Lowe JE, Olsson SB, Prystowsky EN, Tamargo JL, Wann S, Smith SC Jr, Jacobs AK, Adams CD, Anderson JL, Antman EM, Halperin JL, *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. *Circulation* 2006; **114**: e257–354.

10 Hylek EM, Skates SJ, Sheehan MA, Singer DE. An analysis of the lowest effective intensity of prophylactic anticoagulation for patients with non-rheumatic atrial fibrillation. *N Engl J Med* 1996; **8**: 540–6.

11 Hylek EM, Go AS, Chang Y, Jensvold NG, Henault LE, Selby JV, Singer DE. Effect of intensity of oral anticoagulation on stroke severity and mortality in atrial fibrillation. *N Engl J Med* 2003; **349**: 1019–26.

12 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**: 2049–56.

13 Garcia DA, Regan S, Crowther M, Hylek EM. The risk of hemorrhage among patients with warfarin-associated coagulopathy. *J Am Coll Cardiol* 2006; **47**: 804–8.

14 Cherry DK, Burt CW, Woodwell DA. National Ambulatory Medical Care Survey: 2001 summary. *Adv Data* 2003; **11**: 1–44.

15 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**: 2864–70.

16 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.

17 Kent DL, Vermes D, McDonell M, Henikoff J, Fihn SD. A model for planning optimal follow-up for outpatients on warfarin anticoagulation. Warfarin Optimal Outpatient Follow-up Study Group. *Med Decis Making* 1992; **12**: 132–41.

18 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**: 1155–66.

19 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.

20 Haemostasis and Thrombosis and Task Force for the British Committee for Standards in Haematology. Guidelines on oral anticoagulation: third edition. *Br J Haematol* 1998; **101**: 374–87.

21 Fitzmaurice DA, Murray ET, McCahon D, Holder R, Raftery JP, Hussain S, Sandhar H, Hobbs FD. Self management of oral anticoagulation: randomised trial. *BMJ* 2005; **331**: 1057.

22 Mant J, Hobbs FD, Fletcher K, Roalfe A, Fitzmaurice D, Lip GY, Murray E. 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 randomized controlled trial. *Lancet* 2007; **370**: 493–503.

23 Fihn SD, McDonell MB, Vermes D, Henikoff JG, Martin DC, Callahan CM, Kent DL, White RH. A computerized intervention to improve timing of outpatient follow-up: a multicenter randomized trial in patients treated with warfarin. National Consortium of Anticoagulation Clinics. *J Gen Intern Med* 1994; **9**: 131–9.

24 van Walraven C, Oake N, Wells PS, Forster AJ. Burden of potentially avoidable anticoagulant-associated hemorrhagic and thromboembolic events in the elderly. *Chest* 2007; **131**: 1508–15.

25 Hylek EM, Evans-Molina C, Shea C, Henault LE, Regan S. Major hemorrhage and tolerability of warfarin in the first year of therapy among elderly patients with atrial fibrillation. *Circulation* 2007; **115**: 2689–96.

The Accuracy of Clinician Perceptions of "Usual" Blood Pressure Control

Adam J. Rose, MD, MSc[1,2], Stephanie L. Shimada, PhD[1,3], James A. Rothendler, MD[1,3], Joel I. Reisman, BA[1], Peter A. Glassman, MBBS, MSc[4,5], Dan R. Berlowitz, MD, MPH[1,2,3], and Nancy R. Kressin, PhD[1,2]

[1]Center for Health Quality Outcomes and Economic Research, Bedford VA Medical Center, Bedford, MA, USA; [2]Section of General Internal Medicine, Department of Medicine, Boston University School of Medicine, Boston, MA, USA; [3]Department of Health Policy and Management, Boston University School of Public Health, Boston, MA, USA; [4]Center for the Study of Healthcare Provider Behavior, VA Greater Los Angeles, Los Angeles, CA, USA; [5]David Geffen School of Medicine at UCLA, Los Angeles, CA, USA.

BACKGROUND: The term "clinical inertia" is used to describe the failure to manage a chronic condition aggressively enough to bring it under control. The underlying mechanisms for clinical inertia remain poorly understood.

OBJECTIVE: To describe one potential mechanism for clinical inertia, seen through the lens of clinician responses to a computerized hypertension reminder.

DESIGN: Cohort study.

PARTICIPANTS: A total of 509 hypertensive patients from 2 primary care clinics in urban Veterans Health Administration (VA) Medical Centers. All patients had elevated blood pressure (BP) values that triggered a computerized reminder. Given a set of possible responses to the reminder, clinicians asserted at least once for each patient that medication adjustments were unnecessary because the BP was "usually well controlled".

MEASUREMENTS: Using recent BP values from the electronic medical record, we assessed the accuracy of this assertion.

RESULTS: In most instances (57%), recent BP values were not well controlled, with the systolic BP (56%) much more likely to be elevated than the diastolic BP (13%). Eighteen percent of recent systolic BP values were 160 mmHg or greater.

CONCLUSIONS: When clinicians asserted that the BP was "usually well controlled", objective evidence frequently suggested otherwise. This observation provides insight into one potential mechanism underlying clinical inertia.

KEY WORDS: hypertension; ambulatory care; informatics; quality of care; chronic disease.
J Gen Intern Med 23(2):180–3
DOI: 10.1007/s11606-007-0464-1

Received June 5, 2007
Revised September 20, 2007
Accepted November 12, 2007
Published online November 28, 2007

BACKGROUND

The term "clinical inertia" is used to describe the failure to manage a chronic condition aggressively enough to bring it under control.[1] Numerous studies have shown that more aggressive management improves control of hypertension,[2,3] diabetes,[4,5] and hyperlipidemia.[6] While attempts to reduce clinical inertia have had some success,[7,8] a fuller understanding of the context in which it occurs might help in designing better interventions.

The Veterans Health Administration (VA) has several computerized reminders, which are aimed to assist providers in adhering to guideline-recommended care for common medical conditions.[9–11] One reminder focusing on improving care for hypertension not only prompts clinicians to address uncontrolled hypertension, but also collects data regarding the clinical decision-making process.[12] We used data from two VA primary care clinics to evaluate the concordance of electronically recorded blood pressure (BP) values with claims that the BP is "usually well controlled". In so doing, we used the computerized hypertension reminder as a window into the cognitive processes underlying clinical inertia.

METHODS

Patient Sample

Our sample was drawn from a larger study of VA patients with hypertension, conducted between 1/1/02 and 4/21/04.[13] All patients had diagnoses of hypertension, defined by International Classification of Diseases (ICD-9) codes, on at least 2 occasions in 2001 at the primary care clinics of 2 urban tertiary care VA Medical Centers. There were 3 clinics in the original study[13]; the current analysis includes the two sites that employed the computerized hypertension reminder. These 2 clinics employed many clinicians: 41 clinicians wrote at least 200 prescriptions for antihypertensives at 1 site, and 39 clinicians at the other site. The study was approved by all applicable Institutional Review Boards.

The VA Hypertension Reminder

The VA hypertension reminder was designed to help clinicians provide guideline-concordant care.[12,14] When the most recently recorded BP is 140/90 mmHg or above, including BP values

Table 1. Responses to the Computerized Hypertension Reminder

Response	Number	Percent
Medications adjusted or initiated	942	46
Refuses medication adjustment	50	2
Patient being evaluated/referred for resistant or secondary HTN	37	2
Medication change not warranted due to: Patient usually has well-controlled BP on current therapy	555	27
Medication change not warranted due to: Patient has been non-adherent to current Rx regimen	226	11
Medication change not warranted due to: Patient's current therapy is appropriate based on concomitant risk factors and/or other comorbidities	74	4
Medication change not warranted due to: Patient has had unacceptable side effects from previous attempts at more intensive therapy	12	1
Medication change not warranted due to: Patient has limited life expectancy	6	0.3
Medication change not warranted due to: Other	172	8

Distribution of responses to the computerized hypertension reminder in 2071 separate care episodes for 1,580 patients with elevated blood pressure. The wording is copied verbatim from the text of the reminder; response options related to lifestyle interventions are excluded. Because it was possible to give multiple simultaneous responses, percentages do not add up to 100%.

recorded just before opening the electronic medical record (EMR), the reminder prompts clinicians to adjust the medication, to intervene in some other manner, or to supply a reason for not intervening (Table 1). Recording a repeat BP below 140/90 turns off the reminder until a subsequent BP becomes elevated, but this is not recorded as a reminder response.

Although it is possible to respond to the reminder by recommending lifestyle changes to the patient, we examined only responses relating to increasing medical therapy or justifying a decision not to do so. Specifically, we focused on 1 response: "Medication change not warranted due to: Patient usually has well controlled BP on current therapy." Because some patients had multiple reminder responses during the study, we analyzed the first time a clinician claimed that a patient's BP was "usually well controlled".

Analyses

When clinicians asserted that the BP was "usually well controlled", we examined the most recent BP value before the day of the reminder resolution, using data from the vital signs module of the EMR. To the extent available, we similarly examined the most recent 3 BP values and the final BP value recorded during the study. When there were multiple BP measurements on the same day, we used the value with the lowest systolic blood pressure. We reasoned that most clinicians are prepared to accept the lowest BP measurement on a given day. All analyses were conducted using SAS 9.1 (SAS Corporation, Cary NC).

RESULTS

Table 1 shows the frequency of each possible response to the computerized hypertension reminder. The most common response to the reminder was to adjust the medications (46%); the next most common was to assert that no medication adjustment was needed because the BP was "usually well controlled" (27%). This assertion that the BP was "usually well controlled" was made at least once for 509 patients (32% of 1,580); these 509 patients constituted our study sample. Of these 509 patients, the vast majority (477 [94%]) had at least 1 BP value of 140/90 mmHg or higher recorded in the EMR on the day of the reminder resolution.

The mean age of these 509 patients was 67 years (standard deviation 10.7) with 97% male. Black patients comprised 47% of the sample and White patients the remainder. The sample had a high burden of comorbid illness: 52% had hyperlipidemia, 45% coronary artery disease, 40% diabetes, 20% cerebrovascular disease, 15% renal disease, and 14% congestive heart failure. Only 15% of patients had none of these comorbid conditions, and 28% had 3 or more of them.

Table 2 shows the most recent systolic and diastolic blood pressure (SBP and DBP) values before the day of the reminder resolution. Using a criterion of 140/90 mmHg or greater to represent uncontrolled BP,[14,15] such BPs were uncontrolled in 285 patients (57%). It was much more common for the SBP to be uncontrolled than the DBP.

This phenomenon was not limited to the single BP value before the day of the reminder resolution. For example, among the 468 patients (92% of 509) who had at least 3 BP values before the day of the reminder resolution, 2 of the previous 3 BP values were uncontrolled in 34%, and all 3 were uncontrolled in 26%. BP values after the date of the reminder were similarly likely to be uncontrolled. Three hundred and forty-nine (69% of 509) patients had at least 1 "final" BP value on a later date than the reminder resolution (a median of 184 days later); of those, 56% had an SBP of 140 mmHg or greater, with 13% at 160 mmHg or greater.

Table 2. Blood Pressure Measurement Preceding the Computerized Reminder (*n*=509)

Blood Pressure Measurement	Number of Patients (%)
Systolic Blood Pressure (mm/Hg)	
<140	225 (44)
140–149	120 (24)
150–159	72 (14)
160–179	75 (15)
180+	16 (3)
Diastolic Blood Pressure (mm/Hg)	
<90	444 (87)
90–99	49 (10)
100–109	10 (2)
110+	5 (1)

An elevated blood pressure value activated a computerized hypertension reminder for 509 patients. In response, their clinicians stated that medication adjustments were unnecessary because the blood pressure was "usually well controlled". This table shows, for these 509 patients, the blood pressure value prior to the day of the reminder resolution. One patient had no previous blood pressure values.

DISCUSSION

In response to a computerized BP reminder, clinicians in our study frequently asserted that medication adjustments were not necessary because the BP was "usually well controlled". This claim was often in conflict with recorded vital signs from the EMR, especially with regard to the SBP. Many patients also had uncontrolled BP at the end of the study, suggesting that their BP control did not improve over time. Several possible reasons for the discrepancy between physician responses to a computerized reminder and recorded vital signs are discussed below.

1) *Clinicians may not accurately recall recent BP values.*
   Part of the theory of clinical inertia is that clinicians have unrealistic assessments of recent control.[1] This may take the form of selective recall of controlled BP values as opposed to uncontrolled ones. Berlowitz et al.[2] have documented that patients may have many visits with uncontrolled BP, but relatively few where the therapy is increased. It is possible that the reminder could be redesigned to combat this tendency to misinterpret recent BP trends, to the extent that it exists, by explicitly reminding clinicians of recent BP values.
2) *Clinicians may not be aware of or agree with consensus guidelines regarding BP targets.*
   Two important reasons why clinicians may not accept or observe clinical practice guidelines are unfamiliarity or disagreement with the guideline.[16] Hyman and Pavlik have shown that clinician thresholds for the treatment of hypertension are higher than the consensus guideline of 140/90 mmHg, that many clinicians are not familiar with such guidelines, and that familiarity with the guidelines predicts more aggressive treatment.[17] Oliveria et al.[18] have also demonstrated that clinicians are frequently willing to accept BP (especially SBP) in excess of guideline-recommended target values. In our study, if VA guidelines had categorized an SBP of 149 mmHg as "well controlled", this would have reduced the proportion of patients whose most recent BP was "uncontrolled" by almost half.
3) *Clinicians may have access to data not available from the vital signs module of the EMR.*
   For example, clinicians might have access to the results of home BP monitoring, or might have recorded repeat vital signs in their free-text clinical notes as opposed to the vital signs module of the EMR, from which our BP values were drawn.

Several limitations of our study should be noted. First, we used a population of patients from two urban VA primary care clinics. Our patients were predominantly male and had a high burden of comorbidity, which may affect the generalizability of our results. Second, as noted above, clinicians may have documented additional BP values in their clinic notes, but not in the vital signs package of the EMR. However, Borzecki et al.[19] have shown that the addition of chart reviews to automated data does not affect conclusions in VA studies of hypertension. Finally, clinician response to the reminder was not mandatory; our sample represents only the subset of patients whose clinicians chose to resolve the reminder.

In summary, providers' perceptions that patient BPs are "usually well controlled" are frequently in conflict with objective data from the EMR, especially regarding systolic BP. Future studies should examine whether patient or provider characteristics predict clinician perceptions of BP control independently of recent BP values. It is likely that inaccurate recall of recent BP values and unawareness of or disagreement with clinical practice guidelines are contributing factors. For some clinicians, interventions that target these phenomena may be useful in improving patients' BP control.

Acknowledgments: *This research was supported by a grant from the Department of Veterans Affairs Health Services Research and Development Service (TRH01-038, N. Kressin, PI). Drs. Rose and Shimada are supported by grants from the Veterans Administration Department of Academic Affairs. Dr. Kressin is supported by a Research Career Scientist award from the Department of Veterans Affairs, Health Services Research & Development (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.*

Conflicts of Interest: *Dr. Rothendler reports receiving research support from Sanofi-Aventis. Dr. Berlowitz reports receiving research support from Bristol-Myers Squibb. None of the other authors report any potential conflicts of interest.*

Corresponding Author: *Adam J. Rose, MD, MSc; Center for Health Quality Outcomes and Economic Research, 200 Springs Road, Building 70, Bedford, MA 01730, USA (e-mail: adamrose@bu.edu).*

REFERENCES

1. **Phillips LS, Branch WT, Cook CB, et al.** Clinical inertia. Ann Intern Med 2001;135(9):825–34.
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):1957–63.
3. **Okonofua EC, Simpson KN, Jesri A, Rehman SU, Durkalski VL, Egan BM.** Therapeutic inertia is an impediment to achieving the Healthy People 2010 blood pressure control goals. Hypertension 2006;47(3):345–51.
4. **Ziemer DC, Miller CD, Rhee MK, et al.** Clinical inertia contributes to poor diabetes control in a primary care setting. Diabetes Educ 2005;31 (4):564–71.
5. **el-Kebbi IM, Ziemer DC, Musey VC, Gallina DL, Bernard AM, Phillips LS.** Diabetes in urban African-Americans. IX. Provider adherence to management protocols. Diabetes Care 1997;20(5):698–703.
6. **Becker DM, Raqueno JV, Yook RM, et al.** Nurse-mediated cholesterol management compared with enhanced primary care in siblings of individuals with premature coronary disease. Arch Intern Med 1998;158(14):1533–9.
7. **Lester WT, Grant RW, Barnett GO, Chueh HC.** Randomized controlled trial of an informatics-based intervention to increase statin prescription for secondary prevention of coronary disease. J Gen Intern Med 2006;21 (1):22–9.
8. **Ziemer DC, Doyle JP, Barnes CS, et al.** An intervention to overcome clinical inertia and improve diabetes mellitus control in a primary care setting: improving primary care of African Americans with diabetes (IPCAAD) 8. Arch Intern Med 2006;166(5):507–13.
9. **Doebbeling BN, Chou AF, Tierney WM.** Priorities and strategies for the implementation of integrated informatics and communications technology to improve evidence-based practice. J Gen Intern Med 2006;21(suppl 2): 50–7.
10. **Doebbeling BN, Vaughn TE, McCoy KD, Glassman P.** Informatics implementation in the Veterans Health Administration (VHA) healthcare system to improve quality of care. In: Proceedings of the AMIA Annual Symposium 2006:204–8.
11. **Fung CH, Woods JN, Asch SM, Glassman P, Doebbeling BN.** Variation in implementation and use of computerized clinical reminders in an integrated healthcare system. Am J Manag Care 2004;10(11 Pt 2):878–85.
12. **Glassman P, Volpp B, Walder D, et al.** Developing electronic clinical reminders for improving hypertension management: the approach of the Department of Veterans Affairs. The Journal on Information Technology in Healthcare 2003;1(4):251–65.

13. **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):768–74.
14. Department of Veterans Affairs. Clinical Practice Guidelines: Hypertension. (Accessed October 18, 2006, at http://www.oqp.med.va.gov/cpg.)
15. The sixth report of the Joint National Committee on prevention, detection, evaluation, and treatment of high blood pressure. Arch Intern Med 1997;157(21):2413–46.
16. **Cabana MD, Rand CS, Powe NR, et al.** Why don't physicians follow clinical practice guidelines? A framework for improvement. JAMA 1999;282(15):1458–65.
17. **Hyman DJ, Pavlik VN.** Self-reported hypertension treatment practices among primary care physicians: blood pressure thresholds, drug choices, and the role of guidelines and evidence-based medicine. Arch Intern Med 2000;160(15):2281–6.
18. **Oliveria SA, Lapuerta P, McCarthy BD, L'Italien GJ, Berlowitz DR, Asch SM.** Physician-related barriers to the effective management of uncontrolled hypertension. Arch Intern Med 2002;162(4):413–20.
19. **Borzecki AM, Wong AT, Hickey EC, Ash AS, Berlowitz DR.** Can we use automated data to assess quality of hypertension care? Am J Manag Care 2004;10(7 Pt 2):473–9.

Journal of Substance Abuse Treatment 34 (2008) 464–470

Journal of Substance Abuse Treatment

Brief article

Interpersonal violence exposure and alcohol treatment utilization among medical inpatients with alcohol dependence

Emily F. Rothman, (Sc.D.)[a,b,*], Debbie M. Cheng, (Sc.D.)[c,d], Alison Pedley, (B.A.)[e], Jeffrey H. Samet, (M.D., M.A., M.P.H.)[a,d], Tibor Palfai, (Ph.D.)[b,f], Jane M. Liebschutz, (M.D., M.P.H.)[a,d], Richard Saitz, (M.D., M.P.H.)[b,d,g]

[a]*Department of Social and Behavioral Sciences, Boston University School of Public Health, Boston, MA, USA*
[b]*Youth Alcohol Prevention Center, Boston University School of Public Health, Boston, MA, USA*
[c]*Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA*
[d]*Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Boston Medical Center and Boston University School of Medicine, Boston, MA, USA*
[e]*DM-STAT, Inc., Malden, MA, USA*
[f]*Department of Psychology, College of Arts and Sciences, Boston University, USA*
[g]*Department of Epidemiology, Boston University School of Public Health, USA*

Received 15 May 2007; received in revised form 4 June 2007; accepted 1 July 2007

Abstract

The goal of this study was to examine the association between interpersonal violence exposure and utilization of alcohol treatment after medical hospitalizations among adults with alcohol dependence. We analyzed data collected from a prospective cohort of 238 adults with alcohol dependence who were inpatients in a large urban hospital. Participants who reported interpersonal violence victimization had 1.6 times the odds (adjusted odds ratio = 1.64, 95% confidence interval = 0.92–2.91) of receiving alcohol treatment during the year after hospitalization compared to participants with no violence exposure. Recent (past 3 months) exposure to violence was not more strongly related to receipt of treatment than any lifetime violence exposure. Results suggest that a history of interpersonal violence victimization may be associated with an increased odds of alcohol treatment utilization following a medical hospitalization. Therefore, clinicians should be optimistic about identifying and referring patients who have experienced interpersonal violence to alcohol treatment. Moreover, given the potentially high prevalence of interpersonal violence exposure among inpatient populations at large urban hospitals, alcohol treatment providers should develop methods to address both alcohol dependence and violence recovery.

Keywords: Interpersonal violence; Alcohol dependence; Trauma; Substance abuse treatment

1. Introduction

It is well documented that adult alcohol treatment services utilization is influenced by multiple individual and institutional factors, including the severity of dependence, physical and mental health status, socioeconomic status, ethnicity, gender, cognitive ability, marital and employment status, and the availability of services (Arroyo, Westerberg, & Tonigan, 1998; Edlund, Belin, & Tang, 2006; Green, Polen, Dickinson, Lynch, & Bennett, 2002; Kertesz et al., 2006; McAuliffe & Dunn, 2004; Sakai, Ho, Shore, Risk, & Price, 2005; Satre, Knight, Dickson-Fuhrmann, & Jarvik, 2004). The relationship between trauma exposure and substance dependence is well established (Farley, Golding, Young, Mulligan, & Minkoff, 2004; Ullman, Filipas, Townsend, & Starzynski, 2006), and research demonstrates that the prevalence of lifetime exposure to trauma (e.g., physical or sexual abuse) is high among substance-dependent patients in treatment, with estimates ranging from 37% to 47% for

* Corresponding author. Boston University School of Public Health, Talbot Building, 253 West, 715 Albany Street, Boston, MA 02118, USA. Tel.: +1 617 414 1385.
E-mail address: erothman@bu.edu (E.F. Rothman).

0740-5472/08/$ – see front matter
doi:10.1016/j.jsat.2007.07.006

lifetime physical and/or sexual abuse exposure (Dunn, Ryan, & Dunn, 1994; Easton, Swan, & Sinha, 2000; Pirard, Sharon, Kang, Angarita, & Gastfriend, 2005). However, it is not known whether adult survivors of interpersonal violence victimization who have alcohol dependence are any more or less likely to utilize substance use treatment services or programs (e.g., detoxification, counseling, or self-help) than those with dependence and no exposure to interpersonal violence victimization. Interpersonal violence exposure may be positively associated with treatment and self-help utilization, as trauma exposures including violence have been found to be positively related to increased use of medical services (Liebschutz, Geier, Horton, Chuang, & Samet, 2005; Rosenberg et al., 2000) and correlated with a decreased likelihood of alcohol treatment dropout or under-attendance (Copeland & Hall, 1992; Easton et al., 2000). On the other hand, exposure to interpersonal violence may lead to self-medication with alcohol for trauma-related symptoms and could interfere with help seeking for and receipt of alcohol treatment. It is important for substance abuse treatment professionals to know if survivors of interpersonal violence are as likely or more likely to follow through on clinician referrals to treatment programs or services than other patients who screen positive for dependent drinking. This information could be used to build the case in favor of consistent screening and referrals among general patient populations where the co-occurrence of multiple types of psychological disorders resulting from trauma is prevalent, and patients may otherwise be presumed to be unlikely or incapable of making behavioral changes (e.g., "lost causes"). Substance abuse treatment providers may be able to form new and more effective alliances with hospital-based health care providers who serve populations where experiences of interpersonal violence victimization are prevalent if the likelihood that survivors will follow through on clinician referrals to substance abuse treatment is established. Therefore, the goal of the present analysis was to assess whether lifetime exposure and recent exposure to four specific forms of interpersonal violence were associated with the utilization of alcohol treatment services after hospitalization among medical inpatients identified by screening as having alcohol dependence. To our knowledge, this study is the first to examine the relationship between interpersonal violence exposure and alcohol-dependence-related help seeking among a general patient population.

2. Materials and methods

This study was a secondary analysis conducted using data collected prospectively for a randomized, controlled effectiveness trial of an alcohol screening and brief intervention in a large urban teaching hospital. The study was approved by the Institutional Review Board of the Boston University Medical Center. The results of the trial are described elsewhere (Saitz et al., 2007).

2.1. Participants

Between February 2001 and June 2003, all hospital medicine inpatients ages 18 years and older whose physicians did not decline the research contact were approached by trained research associates and invited to participate in screening (N = 7,824). Patients who were fluent in English or Spanish and gave verbal consent completed a screening interview to determine their eligibility (N = 5,813). Eligibility criteria included current (past month) drinking of risky amounts (defined for eligibility as >14 standard drinks per week or ≥5 drinks per occasion for men and >11 drinks per week or ≥4 drinks per occasion for women and people ≥66 years), a Mini-Mental State Examination score of ≥21, consent to participate in two follow-up contacts associated with the study, and no plans to move away from the area during the upcoming year. Of the 986 identified eligible inpatients drinking risky amounts, 341 provided written informed consent and enrolled in the clinical trial (35%). Because only people with alcohol dependence would be expected to utilize alcohol treatment programs or helping services, the current analyses are restricted to that sample.

2.2. Procedure

Eligible participants were interviewed at three time points over a 1-year period, at baseline, 3 months following baseline, and 12 months following baseline. Interviews took place in-person, with the exception of 11% of the 3-month and 13% of the 12-month follow-ups that took place by telephone.

2.3. Measures

2.3.1. Dependent variable

To assess alcohol treatment utilization, at the 3- and 12-month follow-up assessments, patients were read a series of six questions. These six questions elicited information about patients' use of various forms of alcohol and drug assistance, as well as mental and physical health care services, in the past 3 and 12 months, respectively. For the present analysis, we defined "treatment" broadly, beyond formal specialty treatment, and included in our definition forms of assistance such as hospital detoxification; all other detoxification programs; residential alcohol or drug treatment; living in a halfway house; counseling or therapy; attendance at self-help, mutual-help, or 12-step meetings (e.g., Alcoholics Anonymous); or counseling from an Employee Assistance Program. The definition excluded the brief intervention provided by the clinical trial. A positive response to any of the alcohol treatment questions at either the 3- or 12-month assessment was classified as receipt of alcohol treatment during the year following medical hospitalization.

2.3.2. Independent variable

Our primary independent variable of interest, exposure to interpersonal violence, was a binary variable created from four separate questions asked at baseline that solicited information about lifetime and recent experiences with (a) family or partner physical violence, (b) stranger-perpetrated violence (i.e., robbing, mugging, or physical attacks), (c) sexual assault, and (d) rape. These four questions were selected from the 31 items used to assess trauma exposure in the Women, Co-occurring Disorders, and Violence Study (McHugo et al., 2005) and are modified versions of items from the Life Stressor Checklist-Revised (Wolfe & Kimerling, 1997). The content validity of these questions has been determined to be good, and the reliability of these four questions has been assessed in a sample of women with substance abuse disorders and has been found to be adequate (κ values = .52–.63; McHugo et al., 2005). Individuals who reported that they had experienced any one of the four interpersonal violence victimization exposures at any point during their lives were classified as having lifetime interpersonal violence exposure. These individuals were asked follow-up questions to determine when these experiences had occurred. For example, they were asked, "How old were you when this first happened?" and to distinguish the effects of recent (defined as occurring in the past 3 months) and lifetime experiences of trauma, we also categorized subjects into three groups: any interpersonal violence exposure within the past 3 months (i.e., recent exposure), any lifetime exposure but no recent exposure, and never exposed.

Covariates included in the analyses were randomization group, age, gender, White race versus other race, homelessness and employment status, posttraumatic stress disorder (PTSD) symptoms, and the frequency and number of alcohol-related consequences or problems. All covariates were assessed at baseline, self-reported and assessed through single questions, with the exception of PTSD symptoms and severity of alcohol dependence. PTSD was assessed via the civilian PTSD symptom checklist (PCL-C), which is a valid, reliable 17-item self-report symptom scale whose score corresponds to the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* diagnosis of PTSD (Blake et al., 1995; Blanchard, Jones-Alexander, Buckley, & Forneris, 1996). The severity of alcohol consequences or problems was measured via the Short Inventory of Problems (SIP) total score (Miller, Tonigan, & Longabaugh, 1995). Time since baseline was also included as a covariate in all adjusted regression analyses.

2.4. Data analysis

Study data were analyzed using SAS/STAT software versions 8.2 and 9.1.3 (SAS Institute, Inc., Cary, NC). Frequencies and proportions were used to describe the study sample. Bivariate associations between baseline characteristics and treatment utilization were assessed using a chi-square test. We used generalized estimating equations (GEE) with a logit link function to assess the association between trauma and treatment utilization, adjusting for randomization group, age, gender, race, homelessness, and employment status. The GEE approach was used to account for the correlation due to repeated measures from the same subject over time. The empirical standard errors from the GEE approach were used for all analyses. Models did not include variables that were highly correlated with each other (correlations >.40) to minimize the potential for collinearity. PTSD symptoms (PCL-C score) and severity of alcohol problems (SIP score) were not included as covariates in the primary regression models because these factors were expected to be in the causal pathway between interpersonal violence exposure and treatment utilization rather than confounders. To further assess this relationship, we fit secondary models that included these variables as covariates to assess whether their inclusion strongly attenuated the effect of violence exposure. PTSD symptoms and severity of alcohol problems were included separately in models as they were found to be highly correlated (correlation coefficient = .55). All analyses were conducted with two-sided tests and a significance level of .05.

3. Results

Of the 341 subjects enrolled in the randomized clinical trial, 261 (77%) had alcohol dependence. Of these, 238

Table 1
Characteristics of medical inpatients with alcohol dependence, by lifetime exposure to interpersonal violence victimization (N = 238)

Characteristics	Total (N = 238), n (%)	With violence exposure (n = 183), n (%)	Without violence exposure (n = 55), n (%)	p
Gender				.0186
Male	169 (71.0)	123 (67.2)	46 (83.6)	
Female	69 (29.0)	60 (32.8)	9 (16.4)	
Ethnicity				.7865
Hispanic	20 (8.4)	15 (8.2)	5 (9.1)	
Non-Hispanic	218 (91.6)	168 (91.8)	50 (90.9)	
Race				.8643
White	85 (35.7)	68 (37.2)	17 (30.9)	
Black	116 (48.7)	87 (47.5)	29 (52.7)	
Employed (past 3 months)				.8527
Yes	76 (31.9)	59 (32.2)	17 (30.9)	
No	162 (68.1)	124 (67.8)	38 (69.1)	
Homeless (≥1 night, past 3 months)				.0591
Yes	72 (30.3)	61 (33.3)	11 (20.0)	
No	166 (69.7)	122 (66.7)	44 (80.0)	
Age (years)				.1466
18–25	11 (4.6)	9 (4.9)	2 (3.6)	
26–35	32 (13.4)	28 (15.3)	4 (7.3)	
36–45	91 (38.2)	73 (39.9)	18 (32.7)	
≥46	104 (43.7)	73 (39.9)	31 (56.4)	

Note. Data are from 238 patients, at baseline.

(91%) had follow-up data on treatment utilization at the 3- or 12-month time points and comprised the final sample for analysis. Subjects ($N = 238$) were primarily male (71%), and most subjects were not Hispanic (92%) and unemployed (68%; Table 1). Thirty percent of the sample reported that they were homeless, which means having had at least one night in a shelter or street in the past 6 months.

There was an exceedingly high prevalence of interpersonal violence victimization among this sample. At the time of hospitalization (i.e., baseline), 77% of the inpatients with alcohol dependence reported at least one experience of interpersonal violence victimization in their lifetimes. Of those who had been victimized, 46% reported family or partner physical abuse victimization, 57% reported experiencing stranger-perpetrated violence, 29% reported having been sexually assaulted, and 22% reported a history of rape. The risk for violence exposure was slightly elevated among females (relative risk = 1.19, not shown).

One year after hospitalization, 56% of all alcohol-dependent patients identified through clinician screening reported having used any alcohol treatment program or service in the past 12 months. The unadjusted analysis revealed that participants who reported any lifetime history of interpersonal violence victimization, including physical abuse, stranger-perpetrated violence, sexual assault, or rape, had almost twice the odds (odds ratio [OR] = 1.88, 95% confidence interval [CI] = 1.08–3.28) of treatment utilization in the year following medical hospitalization compared to participants with no lifetime history of violence exposure (Table 2). When specific types of interpersonal violence were assessed individually, a higher odds of treatment utilization was consistently observed for those who had experienced each of the four forms of violence as compared to those who had not, and the effect was strongest for rape (OR = 2.23, 95% CI = 1.23–4.05; Table 2). A similar effect was observed in the adjusted models. In the multivariable regression model adjusting for randomization group, age, gender, race, homelessness, and employment status, the odds of treatment utilization were higher for those with any lifetime history of violence exposure compared to those with no such exposure (OR = 1.64, 95% CI = 0.92–2.91), although the difference was not statistically significant at the $p < .05$ level (Table 2). A lifetime history of being a victim of stranger violence remained significantly related to treatment utilization in the adjusted analysis (OR = 1.70, 95% CI = 1.03–2.80); however, the effect of the other forms of violence victimization was attenuated and no longer statistically significant. To explore our post hoc hypotheses that PTSD, the severity of drinking problems, or both were potential mediators of the relationship between any lifetime violence exposure and treatment

Table 2
Use of alcohol treatment services by medical inpatients with alcohol dependence during the year following hospitalization, by lifetime exposure to interpersonal violence ($n = 230$)

Type of interpersonal violence	*n*	Use of alcohol treatment services (%)	Unadjusted OR (95% CI)	Adjusted OR (95% CI) [a]
Any violence (composite)				
Yes	188	57	1.88 (1.08–3.28) *	1.64 (0.92–2.91)
No	42	42		
Physical abuse [b]				
Yes	118	61	1.72 (1.07–2.76) *	1.48 (0.89–2.45)
No	109	47		
Stranger-perpetrated violence [c]				
Yes	144	60	1.68 (1.05–2.70) *	1.70 (1.03–2.80) *
No	86	46		
Sexual assault or rape (composite)				
Yes	87	64	1.87 (1.12–3.12) *	1.54 (0.88–2.68)
No	143	49		
Sexual assault [d]				
Yes	78	64	1.81 (1.06–3.09) *	1.46 (0.84–2.55)
No	152	49		
Rape [e]				
Yes	67	69	2.23 (1.23–4.05) **	1.67 (0.85–3.28)
No	163	49		

Note. Analyses are based on data from 230 subjects and 430 observations across the 3- and 12-month assessments.

[a] Controlling for randomization group, age, gender, White race, time since baseline, homelessness, and employment.

[b] Response to "Have you ever been physically abused—for example, hit, choked, burned, or beaten—or severely punished—for example, locked up, shut in a closet, tied up, or chained—by someone you knew well such as a parent, sibling, boyfriend, or girlfriend?"

[c] Response to "Have you ever been robbed, mugged, or physically—not sexually—attacked by a stranger or someone you did not know well?"

[d] Response to "Have you ever been touched or made to touch someone else in a sexual way because you felt forced in some way or threatened by harm to yourself or someone else?"

[e] Response to "Have you ever had sex because you felt forced in some way or threatened by harm to yourself or someone else?"

* $p < .05$.

** $p < .01$.

Table 3
Use of alcohol treatment services by medical inpatients with alcohol dependence during the year following hospitalization, by recent and lifetime interpersonal violence exposure (n = 230)

Type of interpersonal violence	n	Use of alcohol treatment services (%)	Unadjusted OR (95% CI)	Adjusted OR (95% CI)[a]
Any interpersonal violence (composite)				
In the past 3 months	30	59	2.11 (0.90–4.95)	1.38 (0.56–3.45)
Any lifetime exposure, but not in the past 3 months	158	57	1.84 (1.04–3.24)*	1.68 (0.94–3.02)
Never	42	42	1.00 (referent)	1.00 (referent)
Physical abuse				
In the past 3 months	14	61	1.73 (0.54–5.51)	1.12 (0.34–3.72)
Any lifetime exposure, but not in the past 3 months	104	60	1.71 (1.05–2.80)*	1.50 (0.90–2.54)
Never	109	47	1.00 (referent)	1.00 (referent)
Stranger-perpetrated violence				
In the past 3 months	16	57	1.69 (0.65–4.38)	1.15 (0.41–3.24)
Any lifetime exposure, but not in the past 3 months	128	60	1.68 (1.03–2.74)*	1.80 (1.07–3.02)*
Never	86	46	1.00 (referent)	1.00 (referent)
Sexual assault				
In the past 3 months	7	64	1.52 (0.34–6.69)	1.13 (0.25–5.19)
Any lifetime exposure, but not in the past 3 months	69	63	1.80 (1.03–3.13)*	1.47 (0.82–2.64)
Never	152	49	1.00 (referent)	1.00 (referent)
Rape				
In the past 3 months	13	76	3.08 (0.81–11.73)	2.19 (0.56–8.55)
Any lifetime exposure, but not in the past 3 months	54	68	2.08 (1.10–3.96)*	1.56 (0.75–3.28)
Never	163	49	1.00 (referent)	1.00 (referent)

Note. Analyses are based on data from 230 subjects and 430 observations across the 3- and 12-month assessments.

[a] Controlling for randomization group, age, gender, White race, time since baseline, homelessness, and employment.

* $p < .05$.

utilization, we fit additional regression models that included each of these variables separately. Both PTSD symptoms and drinking problem severity attenuated the OR corresponding to the primary predictor of interest (violence exposure) by more than 10%, as would be expected if they were mediators of the relationship between violence exposure and treatment utilization.

In adjusted analyses assessing the impact of recent interpersonal violence victimization on treatment utilization, ORs for all categories of interpersonal violence were >1, although none reached statistical significance (any interpersonal violence OR = 1.4, physical abuse OR = 1.1, stranger-perpetrated violence OR = 1.2, sexual assault OR = 1.1, rape OR = 2.2; Table 3). Only recent exposure to rape produced a clinically important increase in the odds of treatment utilization (Table 3). ORs were larger for lifetime-but-not-recent interpersonal violence for all categories except rape (recent rape OR = 2.2, lifetime-but-not-recent rape OR = 1.6; Table 3).

4. Discussion

Prior research indicates that trauma survivors are more likely than people who have not experienced trauma to receive medical services and are less likely to drop out of substance use treatment (Copeland & Hall, 1992; Liebschutz et al., 2005; Rosenberg et al., 2000). Consistent with these studies, we observed an association, with moderate effect sizes, in the relationship between interpersonal violence victimization and subsequent utilization of alcohol treatment services after hospitalization among medical inpatients with alcohol dependence. In addition, these findings were consistent with a mediating effect of both PTSD symptoms and alcohol problem severity, which other research has also detected (Breslau, Davis, & Schultz, 2003).

There are at least three possible explanations for the association between interpersonal violence victimization and increased alcohol treatment utilization. First, it is plausible that the violence victimization preceded PTSD, that PTSD influenced drinking problem severity among survivors, and that the severity of alcohol dependence and consequences brought these individuals into more frequent contact with health professionals and services than individuals who were never victimized. Alternately, patients who experience violence and suffer from PTSD may, as a result, experience problems in multiple areas of their lives simultaneously, which may bring them into contact with providers more frequently than individuals in the general population. This disproportionate contact with providers may result in a greater number of referrals to substance abuse treatment. Finally, evidence suggests that assessment for an alcohol

problem may prompt all alcohol-dependent patients to reflect on their alcohol use, seek help, and decrease consumption (Kypri, Langley, Saunders, & Cashell-Smith, 2007). Patients who have experienced an interpersonal violence trauma during their lives may be particularly primed to react to an assessment, as compared to other alcohol-dependent patients.

This comparison of alcohol treatment utilization among those with recent (past 3 months), lifetime, and no violence victimization exposure suggested that for most types of interpersonal violence experiences, lifetime rather than recent events may be stronger predictors of alcohol treatment utilization in the year following a medical hospitalization, with the exception of rape. In this sample, individuals who had been raped within the 3 months prior to hospitalization had higher odds of using alcohol treatment at the 12-month follow-up than individuals who had been raped during their lifetimes (but not within the prior 3 months). These findings raise several questions for additional research studies with larger samples. For example, it would be beneficial to understand whether there is a "critical window" for alcohol-dependent rape survivors in the immediate aftermath of the trauma that is not necessarily present for other survivors of violence (e.g., survivors of muggings, sexual molestation, or physical assault) and whether rape functions as a cue to help seeking among alcohol-dependent individuals in a way that other forms of violence victimization do not.

Clinicians should recognize that hospitalization may represent a particular opportunity to begin addressing alcohol dependence. Substance abuse treatment professionals can help tailor treatment services so that they address both trauma and alcohol dependence simultaneously. Both substance abuse treatment and health care providers should recognize that patients may be more receptive to substance abuse treatment than treatment for their traumatic experiences because they may feel less stigma, shame, and guilt about being alcohol dependent than having experienced the abuse or violence (Gibson & Leitenberg, 2001; Street, Gibson, & Holohan, 2005); they may not recall the traumatic event (Halligan, Clark, & Ehlers, 2002); or they do not consider the interpersonal violence to be an important determinant of their own mental health (Lab & Moore, 2005).

More research is needed to clarify why individuals reporting no exposure to interpersonal violence may be less likely to initiate and utilize treatment services than those who are survivors of violence and whether additional motivational intervention techniques should be developed to increase treatment utilization among the former subgroup. Moreover, education of clinicians in substance use treatment and medicine should emphasize that trauma survivors are not at all "hopeless" or "lost causes" in terms of their capacity to follow through on referrals to treatment. On the contrary, despite the multiple challenges that they face, these results suggest that survivors of interpersonal violence victimization may be more likely than other patients to seek help and utilize alcohol treatment services pursuant to a hospital-based alcohol screening.

The results of this study are subject to three main limitations. First, treatment utilization data were self-reported. It is possible that violence survivors were more likely to recall and report past month treatment utilization than patients who had not experienced interpersonal violence, although the likelihood of differential reporting is small. Moreover, our definition of "treatment" might not have captured all forms of alcohol treatment. However, a strength of our study is that we used a broad definition of treatment, including common forms of treatment not typically captured in treatment databases (e.g., mutual and self-help groups and Employee Assistance Programs). Second, the sample size was small and may have limited our ability to detect differences in treatment utilization among patients by specific type of violence exposure, particularly in adjusted analyses and analyses of recent violence. Furthermore, although many of the associations in adjusted analyses were not statistically significant, all tested associations (adjusted and unadjusted) had OR point estimates greater than 1. Third, our results may not be generalizable to populations other than inpatients of large academic hospitals in urban settings.

Despite these limitations, our results are encouraging with regard to the likelihood that patients who have faced adversity during their lifetimes, including rape and physical assault, will have the capacity to seek and engage in alcohol dependence treatment and may do so after clinician contact while in the general health care system.

Acknowledgments

This study was supported by the National Institute on Alcohol Abuse and Alcoholism (RO1 12617) and a General Clinical Research Center Grant from the National Center for Research Resources (M01 RR00533).

References

Arroyo, J. A., Westerberg, V. S., & Tonigan, J. S. (1998). Comparison of treatment utilization and outcome for Hispanics and non-Hispanic whites. *Journal of Studies on Alcohol*, *59*, 286−291.

Blake, D., Weathers, F., Nagy, L., Kaloupek, D., Gusman, F., Charney, D., et al. (1995). The development of a clinician-administered PTSD scale. *Journal of Traumatic Stress*, *8*, 75−90.

Blanchard, E., Jones-Alexander, J., Buckley, T., & Forneris, C. (1996). Psychometric properties of the PTSD Checklist (PCL). *Behaviour Research and Therapy*, *34*, 669−673.

Breslau, N., Davis, G., & Schultz, L. (2003). Posttraumatic stress disorder and the incidence of nicotine, alcohol, and other drug disorders in persons who have experienced trauma. *Archives of General Psychiatry*, *60*, 289−294.

Copeland, J., & Hall, W. (1992). A comparison of predictors of treatment drop-out of women seeking drug and alcohol treatment in a specialist women and 2 traditional mixed-sex treatment services. *British Journal of Addiction*, *87*, 883−890.

Dunn, G. E., Ryan, J. J., & Dunn, C. E. (1994). Trauma symptoms in substance-abusers with and without histories of childhood abuse. *Journal of Psychoactive Drugs*, *26*, 357−360.

Easton, C. J., Swan, S., & Sinha, R. (2000). Prevalence of family violence in clients entering substance abuse treatment. *Journal of Substance Abuse Treatment*, *18*, 23−28.

Edlund, M. J., Belin, T. R., & Tang, L. Q. (2006). Geographic variation in alcohol, drug, and mental health services utilization: What are the sources of the variation? *Journal of Mental Health Policy and Economics*, *9*, 123−132.

Farley, M., Golding, J. M., Young, G., Mulligan, M., & Minkoff, J. R. (2004). Trauma history and relapse probability among patients seeking substance abuse treatment. *Journal of Substance Abuse Treatment*, *27*, 161−167.

Gibson, L. E., & Leitenberg, H. (2001). The impact of child sexual abuse and stigma on methods of coping with sexual assault among undergraduate women. *Child Abuse & Neglect*, *25*, 1343−1361.

Green, C. A., Polen, M. R., Dickinson, D. M., Lynch, F. L., & Bennett, M. D. (2002). Gender differences in predictors of initiation, retention, and completion in an HMO-based substance abuse treatment program. *Journal of Substance Abuse Treatment*, *23*, 285−295.

Halligan, S. L., Clark, D. M., & Ehlers, A. (2002). Cognitive processing, memory, and the development of PTSD symptoms: Two experimental analogue studies. *Journal of Behavior Therapy and Experimental Psychiatry*, *33*, 73−89.

Kertesz, S. G., Larson, M. J., Cheng, D. M., Tucker, J. A., Winter, M., Mullins, A., et al. (2006). Need and non-need factors associated with addiction treatment utilization in a cohort of homeless and housed urban poor. *Medical Care*, *44*, 225−233.

Kypri, K., Langley, J. D., Saunders, J. B., & Cashell-Smith, M. L. (2007). Assessment may conceal therapeutic benefit: Findings from a randomized controlled trial for hazardous drinking. *Addiction*, *102*, 62−70.

Lab, D. D., & Moore, E. (2005). Prevalence and denial of sexual abuse in a male psychiatric inpatient population. *Journal of Traumatic Stress*, *18*, 323−330.

Liebschutz, J. M., Geier, J. L., Horton, N. J., Chuang, C. H., & Samet, J. H. (2005). Physical and sexual violence and health care utilization in HIV-infected persons with alcohol problems. *AIDS Care*, *17*, 566−578.

McAuliffe, W. E., & Dunn, R. (2004). Substance abuse treatment needs and access in the USA: Interstate variations. *Addiction*, *99*, 999−1014.

McHugo, G. J., Kammerer, N., Jackson, E. W., Markoff, L. S., Gatz, M., Larson, M. J., et al. (2005). Women, Co-occurring Disorders, and Violence Study: Evaluation design and study population. *Journal of Substance Abuse Treatment*, *28*, 91−107.

Miller, W., Tonigan, J., & Longabaugh, R. (1995). *The Drinker Inventory of Consequences (DrInC): An instrument for assessing adverse consequences of alcohol abuse* (No. NIH Pub No. 95-3911). Washington, DC: Government Printing Office.

Pirard, S., Sharon, E., Kang, S. K., Angarita, G. A., & Gastfriend, D. R. (2005). Prevalence of physical and sexual abuse among substance abuse patients and impact on treatment outcomes. *Drug and Alcohol Dependence*, *78*, 57−64.

Rosenberg, H. J., Rosenberg, S. D., Wolford, G. L., Manganiello, P. D., Brunette, M. F., & Boynton, R. A. (2000). The relationship between trauma, PTSD, and medical utilization in three high risk medical populations. *International Journal of Psychiatry in Medicine*, *30*, 247−259.

Saitz, R., Palfai, T. P., Cheng, D. M., Horton, N. J., Freedner, N., Dukes, K., et al. (2007). Brief intervention for medical inpatients with unhealthy alcohol use—A randomized, controlled trial. *Annals of Internal Medicine*, *146*, 167−176.

Sakai, J. T., Ho, P. M., Shore, J. H., Risk, N. K., & Price, R. K. (2005). Asians in the United States: Substance dependence and use of substance-dependence treatment. *Journal of Substance Abuse Treatment*, *29*, 75−84.

Satre, D. D., Knight, B. G., Dickson-Fuhrmann, E., & Jarvik, L. F. (2004). Substance abuse treatment initiation among older adults in the GET SMART program: Effects of depression and cognitive status. *Aging & Mental Health*, *8*, 346−354.

Street, A. E., Gibson, L. E., & Holohan, D. R. (2005). Impact of childhood traumatic events, trauma-related guilt, and avoidant coping strategies on PTSD symptoms in female survivors of domestic violence. *Journal of Traumatic Stress*, *18*, 245−252.

Ullman, S. E., Filipas, H. H., Townsend, S. A., & Starzynski, L. L. (2006). Correlates of comorbid PTSD and drinking problems among sexual assault survivors. *Addictive Behaviors*, *31*, 128−132.

Wolfe, J., & Kimerling, R. (1997). Gender issues in the assessment of posttraumatic stress disorder. In J. Wilson, & T. Keane (Eds.), *Assessing psychological trauma and PTSD* (pp. 192−238). New York: Guilford.

Relationship of Age of First Drink to Alcohol-Related Consequences Among College Students with Unhealthy Alcohol Use

Emily F. Rothman, ScD
William DeJong, PhD
Tibor Palfai, PhD
Richard Saitz, MD, MPH

ABSTRACT. This study investigated the relationship between age of first drink (AFD) and a broad range of negative alcohol-related outcomes among college students exhibiting unhealthy alcohol use. We conducted an anonymous on-line survey to collect self-report data from first-year college students at a large northeastern university. Among 1,792 respondents who reported ever drinking, 14% reported an AFD before age 14. These early onset drinkers were more likely than later onset drinkers to report frequent drinking, heavy drinking, and other unhealthy alcohol use behaviors. Among the subset of drinkers with unhealthy alcohol use (36%), early drinkers were more likely than later onset drinkers to report experiencing five out of 13 alcohol-related consequences, including driving while intoxicated, missing work or school due to drinking, getting into trouble at work or school due to drinking, receiving lower grades than they should have due to drinking, and developing a tolerance to alcohol. doi:10.1300/J465v29n01_05 *[Article copies available for a fee from The Haworth Document Delivery Service: 1-800-HAWORTH. E-mail address: <docdelivery@haworthpress.com> Website: <http://www.HaworthPress.com>*

KEYWORDS. Age of first drink, unhealthy alcohol use, college health, adolescence

BACKGROUND

Unhealthy alcohol use is commonplace on college campuses in the U.S., and alcohol-related morbidity and mortality among college students is a major public health concern (1). Unhealthy alcohol use ranges from episodically drinking risky amounts to alcohol abuse

Emily F. Rothman and William DeJong are affiliated with Boston University School of Public Health, Boston, MA.
Tibor Palfai is affiliated with the Department of Psychology, Boston University, MA.
Richard Saitz is affiliated with Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Boston Medical Center, Boston University Schools of Medicine and Boston University School of Public Health, Boston, MA.
All authors are affiliated with the Youth Alcohol Prevention Center, Boston University School of Public Health, Boston, MA.
Address correspondence to: Emily F. Rothman, ScD, Assistant Professor, Boston University School of Public Health, Department of Social and Behavioral Sciences, 715 Albany Street, T253W, Boston, MA 02118 (E-mail: EmFaith@aol.com).
This work was supported by grant number 5 P60 AA 013759 from the National Institute on Alcohol Abuse and Alcoholism.

Substance Abuse, Vol. 29(1) 2008
Available online at http://suba.haworthpress.com

doi:10.1300/J465v29n01_05

and alcohol dependence disorders (2). In 2001, over 1,700 college students died of alcohol-related unintentional injuries, while 11% were non-fatally injured because of their drinking (3). As many as 31% of U.S. college students meet diagnostic criteria for alcohol abuse; 6% are alcohol dependent (4).

The prevention of both unhealthy alcohol use and its consequences are important public health objectives (5). Harmful alcohol-related behaviors can significantly compromise the health and safety of both individual drinkers and their communities. Among college students, for example, heavy drinking predicts unsafe sex, physical and sexual assault, criminal violations, physical and cognitive impairment, and interpersonal and academic problems (6,7).

The age of first drink (AFD) is an established risk marker for future alcohol-related problems. AFD is related to high alcohol consumption among high school, college, and adult populations (8-11), plus specific negative alcohol-related consequences such as unintentional injury (12), physical fighting (13), drug use (14), unsafe sex (15), driving after drinking, and accepting rides from intoxicated drivers (16). One study suggests that the relationship between AFD and alcohol dependence holds even among a subset of drinkers with a current pattern of unhealthy alcohol use (17).

Evidence about the potential contribution of AFD to alcohol-related problems among college students with unhealthy alcohol use could inform prevention efforts in two ways: (1) AFD may be a useful marker for identifying high-risk students who need intervention, and (2) existing interventions for college students may be strengthened by addressing age of drinking onset and the alcohol use trajectories of early onset drinkers. To our knowledge, no prior study has investigated AFD as a correlate of alcohol-related consequences among college students with unhealthy alcohol use. The goal of the present study was to investigate the relationship between AFD and a broad range of possible negative alcohol-related outcomes using a sample of first-year students attending a single U.S. university.

METHODS

Sample

In October 2004, all first-year students ages 18 years or older at a large university in the Northeast were sent email invitations to participate in an online health survey. The survey was the first part of an intervention study designed to test the efficacy of a web-based brief intervention to reduce alcohol use; results of the intervention study are available elsewhere (18). The findings reported here are unrelated to the intervention. Study protocols were approved by the Institutional Review Boards at both the study site and Boston University Medical Center.

Procedure

The dean of students informed students about the study via email. The study team sent a second email with a website link to the online survey. Non-respondents received two follow-up reminder emails. All respondents were eligible for prize drawings for $50 or $100 gift certificates or a personal music player. Participant anonymity was maintained by not linking student responses to any identifiers and by directing respondents who completed the survey to a separate website where they could enter their email address for notification if they won a prize.

Questionnaire

We employed a 44-item online survey instrument. In the following order, students were presented with questions about their age and sex, six questions about various health behaviors (e.g., exercise, sleeping, smoking); 13 items about alcohol use, including the 10-item Alcohol Use Disorders Identification Test (AUDIT) (19) and three original alcohol-use questions; four demographic questions (ethnicity, race, height and weight); one question about the age of first drink (AFD); and one question about family history of problem drinking. A subset of respondents who scored 8 or greater on the AUDIT were also presented with 13 questions about alcohol-related conse-

quences from the Young Adult Alcohol Problems Screening Test (YAAPST), described in greater detail below.

All alcohol-related questions are described in greater detail below. Respondents who answered the question "Do you smoke cigarettes?" affirmatively were classified as smokers. Respondents' average number of hours of sleep per night was assessed by asking, "On a typical night during the past month, how many hours did you usually sleep?"

Age of First Drink

Age of first drink (AFD) was assessed through a single item: "About how old were you when you had your first drink of alcohol, other than a few sips?" This item has been used to assess the age of drinking onset on national surveys and has adequate test-retest reliability (kappa = 0.72) (11). Response categories included "I never drank," "21 or over," each age from 12-20 years old, and ages 10-11, 6-9, and 1-5. Those who reported never drinking (n = 402) were removed from the analysis, and the remainder were classified as "ever drinkers." Among ever drinkers, respondents who indicated that they first began drinking at least once a week at age 13 or younger were classified as "early onset drinkers"; those who began to do so at ages 14 or older were classified as "later onset drinkers." Age 14 was selected as a cutoff in order to be consistent with prior studies on AFD and negative alcohol-related consequences (12,17).

Dependent Variables: Alcohol Use and Consequences

Alcohol Use

Survey respondents completed the AUDIT, a 10-item scale frequently used to identify unhealthy alcohol use. Three additional alcohol use questions were also added to this section of the survey: (1) "On how many days do you drink alcohol in a typical week?", (2) "Thinking about the past month, how many times have you had 4 (for women) or 5 (for men) drinks on a given occasion?" and (3) "Thinking about the past month, what are the most drinks you've had on any one occasion?" For AUDIT items, respondents who answered "monthly," "weekly," "daily or almost daily" were compared to those who answered "less than monthly" or "never." The AUDIT has good validity and reliability with both non-college and college populations (20-24). Using a cutoff score of 8 or greater, the AUDIT can correctly identify 97% of people who have an alcohol use disorder (23). Thus, respondents who scored 8 or higher on the AUDIT items were classified as having unhealthy alcohol use, and those who scored less than 8 as non-hazardous drinkers (25).

Negative Alcohol-Related Consequences

Respondents who scored 8 or higher on the AUDIT were directed to complete 13 additional items taken from the Young Adult Alcohol Problems Screening Test (YAAPST). The YAAPST, which also has good validity and internal consistency (26,27), asks respondents to report whether they have experienced each of 26 negative alcohol-related consequences in the past year. We selected 13 of the questions for the survey, excluding items that we anticipated would be either common or exceedingly rare among college students. Examples are: "Have you participated in drinking contests or drinking games (e.g., 'quarters, chugging contests, 'progressive parties')?" and "Has a doctor ever told you that your drinking was harming your health?" Examples of YAAPST items that we used are: "Have you driven a car when you knew you had too much to drink to drive safely?" and "Have you felt very sick to your stomach or thrown up after drinking?" The YAAPST uses multi-point response scales, but for this analysis we collapsed the responses into two categories to indicate presence vs. absence, consistent with some prior studies (e.g., Kahler et al., 2004).

Statistical Analyses

All analyses were carried out using SAS/STAT software version 8.2 (SAS Institute Inc., Cary, NC). First, we compared demographic characteristics (sex and race), AUDIT scores, and two health behaviors (average hours of

sleep per night and smoking) for early and later onset drinkers. Next, among students with unhealthy alcohol use (AUDIT $\geq$ 8), we calculated the proportion of early and later onset drinkers who reported experiencing each of the 13 negative alcohol-related consequences from the YAAPST. Bivariate comparisons were performed with two-sample t-tests for continuous variables and chi-square tests for categorical variables. We fit logistic regression models adjusting for sex and race to examine associations between AFD and the 13 YAAPST items. Two-tailed significance tests were conducted, with alpha set at 0.05.

RESULTS

Characteristics of the Sample

Fifty-five percent of 4,008 eligible first-year students participated in the survey (n = 2,194). Of these, 402 reported never having drunk alcohol and were excluded from the analyses, resulting in a final sample of 1,792 ever drinkers. Characteristics of these students appear in Table 1. The mean age of the respondents was 18 years. Males were slightly underrepresented among the sample compared to the entire first-year class (36% vs. 40%, respectively). The distribution of race and ethnicity among the sample was comparable to that of the first-year class. Approximately 74% of the sample described themselves as White (Non-Hispanic), 2% as Black (Non-Hispanic), 7% as Hispanic, and 18% as another race/ethnicity.

Early Onset of Drinking

Fourteen percent of ever drinkers reported that they had their first drink prior to age 14. More early onset drinkers (55%) than later onset drinkers (34%) were male, and more early onset drinkers (10%) than later onset drinkers (6%) were Hispanic (see Table 1). Early onset drinkers were approximately 1.5 times more likely than later onset drinkers to exhibit unhealthy alcohol use (AUDIT = 8), and 1.2 times more likely to have had 6 or more drinks on one occasion (76% vs. 61%, respectively). Early onset drinkers reported that they got somewhat less sleep per night than later onset drinkers (6.3 hours vs. 6.5 hours, respectively), and they were

TABLE 1. Characteristics of the College Students with Ever Alcohol Use, by Early vs. Later Age of Drinking Onset (n = 1,792)

	All (n=1,792)	Early onset drinkers † (n=258)	Later onset drinkers †† (n=1534)	χ^2 *, p-value
Sex				12.64, p<.001
Male	36%	55%	34%	
Female	64%	45%	66%	
Race				8.31, p<.05
White (non-Hispanic)	74%	75%	73%	
Black (non-Hispanic)	2%	2%	2%	
Hispanic	7%	10%	6%	
Other	18%	13%	19%	
Current drinking behavior				
Unhealthy alcohol use (AUDIT ? 8)	36%	50%	34%	25.98, p<.001
Ever drink 6 or more drinks on occasion	63%	76%	61%	23.42, p<.001
Other health status indicators				
Hours of sleep per night (mean)	6.44	6.29	6.46	1.38, p<.001§
Current smoker	12%	21%	10%	24.72, p<.001

† Began drinking prior to age 14
†† Began drinking at age 14 or older
§ Two-sample independent *t*-test, (df=116) comparing early and later onset drinkers
* Chi-square test comparing early and later onset drinkers

more than twice as likely as later onset drinkers to be smokers (21% vs. 10%, respectively).

Unhealthy Alcohol Use and AFD

Early onset drinkers were substantially more likely than later onset drinkers to report heavy and frequent drinking (see Table 2). Looking at the other alcohol use items, early onset drinkers were more likely to drink frequently (2 or more times per week); consume 7 or more drinks on a typical day when drinking; and drink heavily (6 or more drinks on one occasion) on a daily or weekly basis, controlling for sex and race (see

TABLE 2. Drinking and Alcohol Problems Among College Students Who Report Ever Drinking, by Early vs. Later Age of Drinking Onset (n = 1,792)

AUDIT ITEMS	All (n=1,792)	Early onset drinkers† (n=258)	Later onset drinkers†† (n=1534)	χ^2 *	Adjusted OR (95% CI)§
(1) How often do you have a drink containing alcohol? (2 or more times per week)	34%	45%	32%	p<.0001	1.67 (1.27-2.19)
(2) How many drinks containing alcohol do you have on a typical day when you are drinking? (7 or more drinks)	17%	29%	15%	p<.0001	2.03 (1.47-2.80)
(3) How often do you have six or more drinks on one occasion? (weekly, daily or almost daily)	20%	30%	18%	p<.0001	1.72 (1.26-2.34)
(4) How often during the last year have you found that you were not able to stop drinking once you had started? (monthly, weekly, daily or almost daily)	6%	9%	5%	p<.0001	1.89 (1.15-3.08)
(5) How often during the last year have you failed to do what was normally expected from you because of drinking? (monthly, weekly, daily or almost daily)	4%	8%	4%	p<.01	2.29 (1.35-3.91)
(6) How often during the last year have you needed a first drink in the morning to get yourself going after a heavy drinking session? (any response other than "never")	2%	5%	1%	p<.0001	4.75 (2.23-10.08)
(7) How often during the last year have you had a feeling of guilt or remorse after drinking? (monthly, weekly, daily or almost daily)	5%	5%	5%	NS	1.17 (0.65-2.12)
(8) How often during the last year have you been unable to remember what happened the night before because you had been drinking? (monthly, weekly, daily or almost daily)	9%	15%	8%	p<.01	1.83 (1.23-2.71)
(9) Have you or someone else been injured as a result of your drinking? (ever)	14%	19%	13%	p<.05	1.55 (1.09-2.20)
(10) Has a relative or friend or a doctor or another health worker been concerned about your drinking or suggested you cut down? (ever)	5%	9%	4%	p<.001	2.31 (1.40-3.83)
ADDITIONAL ITEMS					
In a typical week, drinks 4-7 days per week	3%	6%	2%	p<.001	2.91 (1.57-5.38)
Drank heavily 10 or more times in the past month	8%	15%	7%	p<.0001	2.05 (1.38-3.05)
Had 12 or more drinks on one occasion in the past month	12%	19%	11%	p<.0001	1.68 (1.16-2.44)

§ Logistic regression analysis that included sex (male female) and race (Black, Hispanic, other) as control variables
† Began drinking prior to age 14
†† Began drinking at age 14 or older
* Chi-square test comparing early and later onset drinkers

Table 2). Early onset drinkers were also more likely to report drinking 4-7 days per week, drinking heavily 10 or more times in the past month, and having 12 or more drinks on one occasion in the past month.

Negative Consequences Experienced by Students with Unhealthy Alcohol Use

Fully 36% of students who reported ever drinking were classified as having unhealthy alcohol use (AUDIT = 8). Past year negative drinking consequences were commonly reported by this group. For example, almost all of these students reported that they had experienced a hangover (89%) or felt sick or like throwing up as a result of drinking (81%) (see Table 3). Fully 44% reported having sex after drinking that they later regretted, just under one-third had driven while intoxicated during their lifetime, and 25% had missed work or school due to drinking.

Among students with unhealthy alcohol use, AFD was associated with five of the 13 YAAPST negative consequences, controlling for sex and race (see Table 3). Early onset drinkers were more likely than later onset drinkers to have driven while intoxicated (40% vs. 30%, respectively); missed work or school due to drinking (32% vs. 23%, respectively); gotten into trouble at work or school due to drinking (15% vs. 6%, respectively); received lower grades than they should have due to drinking (17% vs. 9%, respectively); and developed a tolerance to alcohol (62% vs. 47%, respectively). Early and later onset drinkers were equally likely to report having experienced hangovers, having felt sick or throwing up due to drinking, having lost friends or had problems with their significant other due to drinking, having neglected work or people for two or more days, having had sex that they later regretted, having experienced "the shakes," or having ever felt that they needed or were dependent on alcohol (Table 3).

DISCUSSION

Among college students with unhealthy alcohol use (AUDIT = 8), age of first drink (AFD) was associated with five out of 13 negative alcohol-related consequences experienced in the past year, including driving while intoxicated, missing work or school due to drinking, getting into trouble at work or school, receiving lower grades due to drinking, and developing a tolerance to alcohol. The alcohol-related consequences that did not appear to be associated with the AFD included two items that were highly prevalent (having experienced a hangover and having felt sick or like throwing up) and four that were relatively rare (having lost friends, having neglected work or people for more than two days, having "the shakes," and having ever felt that they were dependent). Two additional consequences also did not appear to be associated with AFD in this subgroup (having had sex that was later regretted and having problems with a significant other due to drinking).

Consistent with prior research, among ever drinkers, AFD was associated with unhealthy alcohol use and higher frequency and quantity of alcohol consumption (8,12,15). AFD was also associated with other health risks, including getting less sleep and cigarette smoking.

Among ever drinkers, males and individuals reporting Hispanic ethnicity were more likely to report early AFD compared to females and non-Hispanics. While prior studies have found early AFD to be more prevalent among males,8 we are not aware of another study that has reported higher incidence of early AFD among Hispanic college students. It should be noted, however, that our study population included only 125 Hispanic students attending a single university; additional studies using larger samples drawn from multiple colleges are needed to investigate this association.

Our findings suggest that AFD is a strong predictor of alcohol-related consequences even when the analyses are limited to those college students who exhibit unhealthy alcohol use. Additional research is needed to clarify why AFD is associated with negative alcohol consequences. One possibility is that early onset drinkers are more likely to put themselves in risky situations when drinking. Alternatively, early AFD may be part of a general pattern of risk-taking behavior related to sensation-seeking and poor impulse control. Another possibility is that AFD may be a marker for other factors that may also occur during early developmental

TABLE 3. Past Year Consequences of Drinking Among College Students with Unhealthy Alcohol Use‡ by Early vs. Later Age of Drinking Onset (n = 650)

Past Year Consequences (YAAPST ITEMS)	All (n=650)	Early onset drinkers† (n=130)	Later onset drinkers†† (n=520)	χ^2*	Adjusted OR§ (95% CI)
Have you ever gotten into trouble at work or school because of drinking?	8%	15%	6%	p<.001	2.79 (1.54-5.07)
Have you ever received a lower grade on an exam or paper than you should have because of drinking?	10%	17%	9%	p<.01	2.15 (1.24-3.73)
Have you ever found you needed larger amounts of alcohol to feel any effect—or that you could no longer get high or drunk on the amount that used to get you high or drunk?	50%	62%	47%	p<.01	1.91 (1.28-2.84)
Have you ever had "the shakes" after stopping or cutting down on drinking (for example, your hands shake so that your coffee cup rattles in the saucer or you have trouble lighting a cigarette)?	4%	5%	3%	NS	1.83 (0.74-4.55)
Have you not gone to work or missed class at school because of drinking, a hangover, or an illness caused by drinking?	25%	32%	23%	p<.05	1.67 (1.09-2.54)
Have you driven a car when you knew you had too much to drink to drive safely?	32%	40%	30%	p<.05	1.48 (0.99-2.21)
Has drinking ever gotten you into sexual situations that you later regretted?	44%	50%	43%	NS	1.44 (0.98-2.13)
Has your drinking ever created problems between your boyfriend/girlfriend (or spouse) or another near relative?	24%	27%	23%	NS	1.31 (0.84-2.04)
Have you ever lost friends (including boyfriends or girlfriends) because of your drinking?	5%	5%	5%	NS	1.18 (0.50-2.80)
Have you ever felt that you needed alcohol or were dependent on alcohol?	10%	10%	10%	NS	1.08 (0.57-2.06)
Have you ever neglected your obligations, your family, your work, or your schoolwork for two or more days in a row because of your drinking?	11%	11%	11%	NS	1.02 (0.55-1.90)
Have you had a hangover (headache, sick to your stomach) the morning after you have been drinking?	89%	88%	89%	NS	0.92 (0.50-1.69)
Have you felt very sick to your stomach or thrown up after drinking?	81%	80%	81%	NS	0.91 (0.56-1.48)

‡ Scored ≥ 8 on the Alcohol Use Disorders Identification Test (AUDIT)
† Began drinking prior to age 14
†† Began drinking at age 14 or older
* Chi-square test comparing early and later onset drinkers
§ Logistic regression analysis that included sex (male/female) and race (Black, Hispanic, other) as control variables

stages (e.g., witnessing inter-parental violence, experiencing child abuse, or associating with delinquent peers), or other problems that occur early in life (such as conduct or behavioral disorders) that are also related to the negative consequences we investigated. It may also be that duration of heavy alcohol use, independent of AFD, is an important factor for certain alcohol-related consequences.

As other investigators have suggested, interventions that make it more difficult for minors to obtain alcohol may increase the age of drinking onset among some youth, decrease the amount of alcohol that adolescents consume,

and ultimately help reduce hazardous drinking and negative alcohol-related consequences experienced by college students (16). Additional research is also needed to evaluate whether behavioral interventions that delay AFD can prevent problematic drinking behavior and alcohol-related consequences during late adolescence and early adulthood.

LIMITATIONS

Five limitations should be noted. First, our sample was drawn from a single university in the Northeast, and the results may not be generalizeable to college students generally or to young adults who are not in college. That said, the prevalence of unhealthy alcohol use was similar to that seen at many other colleges in the U.S.

Second, our response rate was 55 percent-lower than the ideal of 75% or higher (28), but on a par with response rates for major college health surveys (1,29). It is possible that students with the earliest AFD and those experiencing the most severe alcohol-related consequences were disproportionately underrepresented in the sample. If this were true, however, it would likely have biased our results towards the null, producing an underestimate of the true association between AFD and alcohol-related consequences.

Third, we selected individual items from the YAAPST rather than use the entire measure. Studies of the YAAPST's reliability have been conducted on the instrument as a whole and not on individual items. We were interested only in achieving accurate self-report of specific consequences, however, and did not attempt to use summary consequence scale scores.

Fourth, the study sample was limited to first-year students. Negative alcohol-related consequences may be more prevalent among sophomores, juniors, and seniors. Additional studies in this area should include all four classes.

Fifth, our measure of AFD relied upon participant recall which is prone to error as time since the event increases (30). It should be noted, however, that our study offers an improvement over some prior studies of AFD that utilized samples of adults primarily older than college-age (e.g., Hingson et al., 2001; Hingson et al., 2005). Prospective, longitudinal studies would benefit the field.

In conclusion, our findings show that AFD is associated with several negative alcohol-related consequences among college students with unhealthy alcohol use (AUDIT = 8). These results are consistent with prior research demonstrating that AFD predicts negative alcohol-related consequences among college students generally (15,16). To find that this relationship holds up even among a subgroup of high-risk and problem drinkers is striking. Investigators should seek to understand whether AFD is in and of itself a contributing factor to later adolescent and adult drinking patterns and behavior, or whether it is a marker for other factors that may also occur during early developmental stages. Educators and health behavior interventionists should develop new strategies for reaching youth at high risk for early initiation of alcohol use, and the efficacy of these interventions should be formally evaluated. Regardless of whether AFD is a direct contributor to later unhealthy alcohol use or a risk marker for "upstream" causes, targeting youth at risk for early alcohol initiation for special programs and services may result in desired long-term reductions in unhealthy alcohol use, alcohol dependence, and other alcohol-related consequences.

REFERENCES

1. Wechsler H, Lee JE, Kuo M, Seibring M, Nelson TF, Lee H: Trends in college binge drinking during a period of increased prevention efforts: Findings from 4 Harvard School of Public Health College Alcohol Study Surveys: 1993-2001. J Am Coll Health 2002;50(5): 203-17.

2. Saitz R. Clinical practice. Unhealthy alchol use. New England Journal of Medicine. 352(6):596-607, 2005 Feb 10.

3. Hingson R, Heeren T, Winter M, Wechsler H: Magnitude of alcohol-related mortality and morbidity among US college students ages 18-24: Changes from 1998 to 2001. Annu Rev Public Health 2005;26:259-279.

4. Knight J, Wechsler H, Meichun K, Seibring M, Weitzman E, Schuckit M: Alcohol abuse and dependence among U.S. college students. J Stud Alcohol 2002; 63:263-270.

5. US Department of Health and Human Services. Healthy People 2010. Washington, DC: US Department

of Health and Human Services, January, 2000. www.health.gov/healthypeople/

6. Wechsler H, Davenport A, Dowdall G, Moeykens B, Castillo S: Health and behavioral consequences of binge drinking in college: A national survey of students at 140 campuses. JAMA 1994;272:1672-1677.

7. Perkins H: Surveying the damage: A review of research on consequences of alcohol misuse in college populations. Journal of Studies in Alcohol 2002:91-100.

8. Samson H, Maxwell C, Doyle T: The relation of initial alcohol experiences to current alcohol consumption in a college population. J Stud Alcohol 1989;50: 254-260.

9. Pedersen W, Skrondal A: Alcohol consumption debut: Predictors and consequences. J Stud Alcohol 1998;59:32-42.

10. Flory K, Lynam D, Milich R, Leukefeld C, Clayton R: Early adolescent through young adult alcohol and marijuana use trajectories: Early predictors, young adult outcomes, and predictive utility. Dev Psychopathol 2004;16:193-213.

11. Grant BF, Stinson FS, Harford TC: Age at onset of alcohol use and DSM-IV alcohol abuse and dependence: A 12-year follow-up. J Subst Abuse 2001;13: 493-504.

12. Hingson RW, Heeren T, Jamanka A, Howland J: Age of drinking onset and unintentional injury involvement after drinking. JAMA 2000;284:1527-1533.

13. Hingson R, Heeren T, Zakocs R: Age of drinking onset and involvement in physical fights after drinking. Pediatrics 2001;108:872-877.

14. Lo CC: Timing of drinking initiation: A trend study predicting drug use among high school seniors. J Drug Issues 2000;30:525-554.

15. Hingson R, Heeren T, Winter MR, Wechsler H: Early age of first drunkenness as a factor in college students' unplanned and unprotected sex attributable to drinking. Pediatrics 2003;111:34-41.

16. Hingson R, Heeren T, Zakocs R, Winter M, Wechsler H: Age of first intoxication, heavy drinking, driving after drinking and risk of unintentional injury among US college students. J Stud Alcohol 2003;64: 23-31.

17. Hingson R, Heeren T, Winter M: Age of drinking onset and duration, severity, and number of episodes of alcohol dependence. In Research Society on Alcoholism. 2005. Santa Barbara, CA.

18. Saitz R, Palfai T, Freedner N, Winter M, Macdonald A, Lu J, Ozonoff A, DeJong W, Rosenbloom D: Alcohol screening and brief intervention online for college students. Alcohol Clin Exp Res 2005;29:109A.

19. Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M: Development of the Alcohol Use Disorders Identification test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption-II. Addiction 1993;88:791-804.

20. Clements R: A critical evaluation of several alcohol screening instruments using the CIDI-SAM as a criterion measure. Alcohol Clin Exp Res 1998;22:985-993.

21. Allen JP, Litten RZ, Fertig JB, Babor T: A review of research on the Alcohol Use Disorders Identification Test (AUDIT). Alcohol Clin Exp Res 1997;21:613-619.

22. Reinert DF, Allen JP: The Alcohol Use Disorders Identification Test (AUDIT): A review of recent research. Alcohol Clin Exp Res 2002;26:272-279.

23. Chen CH, Chen WJ, Cheng ATA: New approach to the validity of the alcohol use disorders identification test: Stratum-specific likelihood ratios analysis. Alcohol Clin Exp Res 2005;29:602-608.

24. Aertgeerts B, Buntinx F, Bande-Knops J, Vandermeulen C, Roelants M, Ansoms S, Fevery J: The value of CAGE, CUGE, and AUDIT in screening for alcohol abuse and dependence among college freshmen. Alcohol Clin Exp Res 2000;24:53-57.

25. Kokotailo P, Egan J, Gangnon R, Brown D, Mundt M, Fleming M: Validity of the alcohol use disorders identification test in college students. Alcohol Clin Exp Res 2004;28:914-920.

26. Hurlburt S, Sher K: Assessing alcohol problems in college students. J Am Coll Health 1992;45:134-140.

27. Kahler CW, Strong DR, Read JP, Palfai TP, Wood MD: Mapping the continuum of alcohol problems in college students: a Rasch model analysis. Psychol Addict Behav 2004;18(4):322-323.

28. Fowler F: Survey Research Methods. 3 ed. Applied Social Research Methods Series. Vol. 1 Thousand Oaks, CA: Sage Publications; 2002.

29. DeJong W: The role of mass media campaigns in reducing high-risk drinking among college students. J Stud Alcohol 2002;Supplement 14:182-192.

30. Parra G, O'Neill S, Sher K: Reliability of self-reported age of substance abuse involvement onset. Psychol Addict Behav 2003;17:211-218.

doi:10.1300/J465v29n01_05

gressive statin-based cholesterol-lowering therapy. Marchioli et al[8] found that omega-3 supplements produced similar reductions in mortality in the 11,323 patients of the GISSI-Prevenzione trial, regardless of statin use. Omega-3 fatty acids and statins are thought to improve cardiovascular prognosis through different mechanisms of action. Statins do so largely by lowering LDL-C levels, whereas omega-3 fatty acids have neutral to adverse effects on LDL-C levels.[9] Omega-3 fatty acids confer cardiovascular benefits via enrichment of the cell membranes with DHA and EPA, which increase arrhythmic thresholds, improve arterial health, reduce platelet aggregation, and favorably alter autonomic tone. Statins and omega-3 fatty acids each reduce both triglyceride levels and inflammation and provide additive improvements in these parameters when used in combination.[9]

The question regarding additive cardiovascular benefit in the setting of current guideline-based cholesterol treatment holds true for any therapy introduced and widely adopted before the widespread use of statins. For instance, the benefits of aspirin and β-blockers after myocardial infarction cannot be considered irrelevant today simply because their confirmatory trials largely predated the use of statins.

James H. O'Keefe, MD
John H. Lee, MD
Mid America Heart Institute
Kansas City, MO

Roberto Marchioli, MD
Laboratory of Clinical Epidemiology
Santa Maria Imbaro, Italy

Carl J. Lavie, MD
Ochsner Medical Center
New Orleans, LA

William S. Harris, PhD
Sanford Research/USD and
Sanford School of Medicine of the
University of South Dakota
Sioux Falls

1. Marchioli R, Barzi F, Bomba E, et al, GISSI-Prevenzione Investigators. Early protection against sudden death by n-3 polyunsaturated fatty acids after myocardial infarction: time-course analysis of the results of the Gruppo Italiano per lo Studio della Sopravvivenza nell'Infarto Miocardico (GISSI)-Prevenzione. *Circulation*. 2002;105(16):1897-1903.

2. Mozaffarian D, Gottdiener JS, Siscovick DS. Intake of tuna or other broiled or baked fish versus fried fish and cardiac structure, function, and hemodynamics. *Am J Cardiol*. 2006;97(2):216-222. Epub 2005 Nov 21.

3. Mozaffarian D, Lemaitre RN, Kuller LH, et al. Cardiac benefits of fish consumption may depend on the type of fish meal consumed: the Cardiovascular Health Study. *Circulation*. 2003;107(10):1372-1377.

4. Kelly B, Ikonomou M, Higgs D, Oakes J, Dubetz C. Mercury and other trace elements in farmed and wild salmon from British Columbia, Canada. *Environ Toxicol Chem*. 2008 Jan 22;1. Epub ahead of print.

5. Yokoyama M, Origasa H, Matsuzaki M, et al, Japan EPA Lipid Intervention Study (JELIS) Investigators. Effects of eicosapentaenoic acid on major coronary events in hypercholesterolaemic patients (JELIS): a randomised open-label, blinded endpoint analysis [published correction appears in *Lancet*. 2007;370(9583):220]. *Lancet*. 2007;369:(9567)1090-1098.

6. Burr ML, Fehily AM, Gilbert JF, et al. Effects of changes in fat, fish, and fibre intakes on death and myocardial reinfarction: Diet and Reinfarction Trial (DART). *Lancet*. 1989;2(8666):757-761.

7. Bays HE. Safety considerations with omega-3 fatty acid therapy. *Am J Cardiol*. 2007;99(6A):35C-43C. Epub 2006 Nov.

8. Marchioli R, Marfisi RM, Borrelli G, et al. Efficacy of n-3 polyunsaturated fatty acids according to clinical characteristics of patients with recent myocardial infarction: insights from the GISSI-Prevenzione trial. *J Cardiovasc Med (Hagerstown)*. 2007;8(suppl 1):S34-S37.

9. Davidson MH, Stein EA, Bays HE, et al, COMBination of prescription Omega-3 with Simvastatin (COMBOS) Investigators. Efficacy and tolerability of adding prescription omega-3 fatty acids 4 g/d to simvastatin 40 mg/d in hypertriglyceridemic patients: an 8-week, randomized, double-blind, placebo-controlled study. *Clin Ther*. 2007;29(7):1354-1367.

Poor Care, Not Poor Protocols, for Alcohol Withdrawal

To the Editor: I applaud Hecksel et al[1] for their study that found that many general hospital patients were treated inappropriately for alcohol withdrawal syndrome (AWS). Some patients had no recent alcohol use, and others who could not communicate were treated with a strategy that requires communication. However, the authors concluded that symptom-triggered therapy (STT) may be inappropriate in medically or surgically unstable patients or those with a history of alcohol dependence. These conclusions (echoed in the article title) cannot be drawn from their study. Because AWS occurs only in people with alcohol dependence, surely that diagnosis cannot preclude the use of STT. Furthermore, although alcohol symptom scales and STT have been studied less in general hospital patients than in patients in more specialized settings, I agree with the editorialists who concluded that STT was not inappropriate ("we doubt that the protocol itself is at fault"), rather that care was being implemented inappropriately.[2]

There are 2 ways to treat AWS: STT or medications administered on a fixed schedule (FS).[3] Administering medications on an FS risks overdosing and underdosing but is often used to ensure that patients receive at least some benzodiazepine, often in settings in which close monitoring is impossible. However, FS doses should be supplemented (or withheld) on the basis of clinical status (ie, symptoms). Doses of STT, which are delivered on the basis of symptoms, can be implemented after an initial dose for an asymptomatic patient at high risk of complications. These 2 treatments are the only ones recommended on the basis of randomized trial results; no alternatives exist. Given the reality that people with AWS should be treated no matter their location, including in general hospitals, treatment should be with STT or FS strategies. Both approaches have limitations in general hospital patients, but the need for patients to be "monitored more closely" or the requirement of "higher benzodiazepine doses" are both issues that arise regardless of strategy. Neither proven approach limits medication doses, and both should include frequent monitoring (similar to blood sugar management in general hospitals).

The editorialists suggested that AWS should be treated by specialists, in part because protocols can fail in the hands of

physicians from diverse specialties. I hope they are incorrect because, if they are correct, general hospital care would need to come to a screeching halt. About 13% of general medical inpatients have current alcohol dependence.[4] Specialty involvement in every case of AWS is not feasible and is counter to goals of care integration for these patients whose care is so often fragmented. Generalist physicians must implement high-quality care in hospitals for numerous common conditions, including AWS. Evidence shows that they can do this (eg, venous thrombosis prophylaxis).[5] More importantly, they must. We should not conclude that generalist hospital physicians cannot appropriately implement AWS care, although I agree with the authors that it has not yet been proved possible. Evidence-based protocols, along with training and systems that support good-quality care, will likely be required to achieve optimal management of AWS. Although AWS protocols may benefit from improvement, our patients will gain more if we focus on better implementation of known effective care.

Richard Saitz, MD, MPH
Boston University, Boston Medical Center
Boston, MA

1. Hecksel KA, Bostwick JM, Jaeger TM, Cha S. Inappropriate use of symptom-triggered therapy for alcohol withdrawal in the general hospital. *Mayo Clin Proc.* 2008;83(3):274-279.
2. Berge KH, Morse RM. Protocol-driven treatment of alcohol withdrawal in a general hospital: when theory meets practice. *Mayo Clin Proc.* 2008;83(3): 270-271.
3. Mayo-Smith MF, American Society of Addiction Medicine Working Group on Pharmacological Management of Alcohol Withdrawal. Pharmacological management of alcohol withdrawal: a meta-analysis and evidence-based practice guideline. *JAMA.* 1997;278(2):144-151.
4. Saitz R, Freedner N, Palfai TP, Horton NJ, Samet JH. The severity of unhealthy alcohol use in hospitalized medical patients: the spectrum is narrow. *J Gen Intern Med.* 2006;21(4):381-385.
5. Tooher R, Middleton P, Pham C, et al. A systematic review of strategies to improve prophylaxis for venous thromboembolism in hospitals. *Ann Surg.* 2005;241(3):397-415.

To the Editor: The article by Hecksel et al was a welcome and critical review of protocol-driven treatment of alcohol withdrawal. The use of the Revised Clinical Institute for Withdrawal Assessment for Alcohol (CIWA-Ar) protocol was found lacking both in its implementation and in its guidelines for treatment. More than half of the patients who were subjected to the CIWA-Ar protocol were inappropriately selected because they did not meet the inclusion criteria. Moreover, when appropriately selected, the protocol resulted in less than optimal frequency of clinical assessment and inadequate benzodiazepine dosing. Unfortunately, the authors did not address a remedy for such use of the CIWA-Ar protocol. On the basis of their findings, I would not recommend adoption of this protocol in our hospital or in any medical/surgical hospital.

Denis Kennedy, MD
Memorial Health Care System
Chattanooga, TN

To the Editor: The recent study by Hecksel et al highlights a major problem with any protocol for alcohol withdrawal—it can be inappropriately applied to patients who do not actually have alcohol withdrawal. We have certainly seen this occur in our hospital, and strategies for minimizing the inappropriate use of protocols are needed. However, we found aspects of the authors' methods and conclusions problematic.

The authors' criteria for identifying "drinkers" are strict and would likely err in categorizing some patients with substantial risk of developing alcohol withdrawal as "inappropriate." The authors required documentation of recent heavy alcohol consumption (>4 drinks per day for men or >2 drinks per day for women in the week before admission) and a history of alcohol dependence or abuse. However, people with alcoholism are notorious for underrepresenting the amount that they drink and often conceal their drinking from their friends and families, making collateral sources also unreliable. Questions about the quantity and frequency of alcohol intake are known to be insensitive, detecting only 20% to 50% of patients with alcohol abuse or dependence.[1] Beyond the difficulties of quantifying alcohol use, the authors cited no literature to indicate that their thresholds of alcohol intake differentiate patients who are at risk for withdrawal from those who are not. Their criteria also excluded patients who had been drinking heavily if they did not have a previously documented history of alcohol abuse or dependence, the absence of which certainly does not eliminate the risk of withdrawal. Unfortunately, the authors describe all patients not meeting their "heavy consumption" and "history of…abuse" criteria as "nondrinkers," which is misleading and not likely to be true.

The authors required that "patients had to have the ability to communicate meaningfully with nursing staff." They did not discuss how they determined from chart documentation whether patients were communicating "meaningfully." The CIWA-Ar questions are not complicated, so some patients could reasonably answer the yes/no CIWA-Ar questions and still not be able to communicate "meaningfully" in other respects. The authors also specified that patients who were delirious were deemed "inappropriate;" this is confusing since delirium can be part of alcohol withdrawal, and several of the CIWA-Ar subscores grade the severity of symptoms of delirium (such as orientation and clouding of sensorium, agitation, tactile disturbances, auditory disturbances, and visual disturbances).

In their discussion, the authors write, "In general hospitals, patients with a history of alcohol dependence are the most likely to experience adverse events, such as DT [delirium tremens] and withdrawal seizures, when receiving STT according to the CIWA-Ar protocol." This statement implies that patients develop delirium tremens and seizures *because* they are being treated with a CIWA-Ar protocol, which is certainly not what the authors studied (for example, the authors did not compare patients treated with a CIWA-Ar protocol with those managed with another strategy). The authors go on to write "Such patients should be monitored more closely and treated more liberally with benzodiazepines than the

The Case for Chronic Disease Management for Addiction

Richard Saitz, MD, MPH, Mary Jo Larson, PhD, MPA, Colleen LaBelle, RN, Jessica Richardson, MPH, and Jeffrey H. Samet, MD, MA, MPH

Abstract: Chronic disease (care) management (CDM) is a patient-centered model of care that involves longitudinal care delivery; integrated, and coordinated primary medical and specialty care; patient and clinician education; explicit evidence-based care plans; and expert care availability. The model, incorporating mental health and specialty addiction care, holds promise for improving care for patients with substance dependence who often receive no care or fragmented ineffective care. We describe a CDM model for substance dependence and discuss a conceptual framework, the extensive current evidence for component elements, and a promising strategy to reorganize primary and specialty health care to facilitate access for people with substance dependence. The CDM model goes beyond integrated case management by a professional, colocation of services, and integrated medical and addiction care—elements that individually can improve outcomes. Supporting evidence is presented that: 1) substance dependence is a chronic disease requiring longitudinal care, although most patients with addictions receive no treatment (eg, detoxification only) or short-term interventions, and 2) for other chronic diseases requiring longitudinal care (eg, diabetes, congestive heart failure), CDM has been proven effective.

Key Words: chronic disease management, addiction, primary care, linkage, addiction treatment, chronic care model, recovery

(*J Addict Med* 2008;2: 55–65)

Substance (alcohol and drug) dependence is a chronic disease for which many affected adults receive no intervention or detoxification without subsequent treatment. Like other chronic diseases (eg, diabetes, congestive heart failure), substance dependence has no cure and is characterized by relapses requiring longitudinal care. Medical and psychiatric comorbidities are the rule rather than the exception. As a result, care delivery can be complex both for clinicians and patients. In the United States, systems of care for substance dependence (both alcohol and drug) are rarely integrated with those for medical and psychiatric illnesses. Specialty alcohol and drug treatment is efficacious, but many patients do not access available treatment for substance dependence problems after detoxification or medical care. Others enter specialty addiction treatment but do not receive medical or psychiatric care. Some have called to expand the frame of health services research on addictions to include services outside the specialty treatment sector, including behavioral care integrated into primary care.[1–3]

Although primary care settings provide the venue for longitudinal, comprehensive, and coordinated care, their potential to effectively treat addiction and related comorbidities has not been realized; medical, mental health, and addiction treatment are not coordinated. Primary care settings, with reorganization and appropriate service elements, hold the promise of simultaneously improving physical and psychiatric health while decreasing substance dependence problems. Although chronic disease management—longitudinal care delivery linking, integrating, and coordinating primary and specialty health care—is effective for other diseases, it has not been described or studied for substance dependence.

CHRONIC DISEASE (CARE) MANAGEMENT

Chronic disease (care) management (CDM) is based on a chronic care model, as described by Wagner et al.[4–6] Chronic disease (care) management is a patient-centered model of care, which includes patient and clinician education, explicit evidence-based care plans, and expert care availability (Table 1). In using the term "disease management," this

From the Clinical Addiction Research and Education (CARE) Unit (RS, JR, JHS), Boston Medical Center, Boston, MA; Section of General Internal Medicine, Department of Medicine (RS, CL, JHS), Boston Medical Center and Boston University School of Medicine, Boston, MA; Department of Epidemiology, and the Youth Alcohol Prevention Center (RS), Boston University School of Public Health, Boston, MA; New England Research Institutes, Inc. (MJL), Watertown, MA; Department of Social and Behavioral Sciences (JHS), Boston University School of Public Health, Boston, MA.

Ms. Richardson is now at the Division of General Internal Medicine, Albert Einstein College of Medicine, Bronx, NY.

Received May 20, 2007; revised December 28, 2007; accepted December 28, 2007.

Send correspondence and reprint requests to Richard Saitz, MD, MPH, Section of General Internal Medicine, Boston Medical Center, 801 Massachusetts Avenue, 2nd Floor, Boston, MA 02118; e-mail: rsaitz@bu.edu

Drs. Saitz, Samet, and Larson were supported in this work by grants from the National Institute on Alcohol Abuse and Alcoholism (NIAAA) (R-01 AA 10870) and the National Institute on Drug Abuse (NIDA) (R-01 DA 10019). In addition, Dr. Saitz is supported by grants from the NIAAA (R-01s AA. 13216, AA 14258, AA 14713, R25, AA 13822, and P60 AA 13759), the Substance Abuse and Mental Health Services Administration (SAMHSA) UT 9 TI018311, and the National Center for Research Resources (NCRR) (K30 RR 22252). Dr. Samet is supported by grants from NIAAA (R-01 AA 016059, K-24 AA 015674, and R-21 AA 014821), NIDA (R-01 DA 019841 and R-25 DA 13582), and (SAMHSA) (1 U79T8 018311), and Dr. Larson is supported by a grant from the NIDA (R01 DA 016929).

This work was presented in part at the annual meeting of the Research Society on Alcoholism, July 7–12, 2007, Chicago, IL.

ISSN: 1932-0620/08/0202-0055

TABLE 1. Chronic Disease Management (CDM) Conceptual Elements and Potential Elements for Implementation

CDM Conceptual Element	Potential Implementation Elements
Community resources	Case management; address social, legal, financial needs
Chronic disease as priority	Focus on substance dependence as a chronic illness; explicit care plans
Self-management support	Routine assessment and feedback; patient participation (collaborative care); behavior change; psychosocial support
Delivery system design	On-site service delivery (integrated and coordinatedcare); referral agreements; planned visits; use of nonphysicians in multidisciplinary team; patient reminders; collaboration of addiction, medical and psychiatric physicians
Decision support	Specialty expertise made accessible
Clinical information systems	Electronic medical record; patient registry; monitoring of outcomes

article is about health services based on Wagner's model and *not* about corporate programs that focus primarily on patient self-care and are implemented via telephonic contacts.[7]

CDM CONCEPTUAL FRAMEWORK

Many, including the Institute of Medicine (IOM), have recognized the challenge of managing chronic conditions in a health care system designed to treat acute illness.[8] In 1996, Wagner et al. proposed a solution—CDM[4–6]—based on a chronic care model. Chronic disease management is implemented by the following multidisciplinary team members: nurse clinical care managers with disease-specific skills to coordinate referrals, communicate with clinical caregivers, and proactively follow patients; social workers to access community resources; and physician specialists. The U.S. Center for Medicare and Medicaid Services (CMS) and the National Committee for Quality Assurance (NCQA) have recognized the potential of CDM.[9–11] Policy analysts have called for CDM to be integrated into primary care settings to avoid fragmented care.[12–15]

We argue that chronic disease management could be applied to addiction for patients in the primary care setting, particularly those who are out-of-treatment.[6] Of U.S. adults, 23 million have substance dependence, costing society almost $300 billion each year, yet 90% of people with addictions are out-of-treatment because of individual barriers and a fragmented treatment system.[16–23] Thus, although addictions treatment is effective and reduces health care costs, its benefits are largely unrealized.

In addition to the chronic care model, 2 additional conceptual models support the categories and elements specified in a CDM approach applied to addictions: D'Aunno[24] and others' integrative linkages of health services, and Andersen's[25] behavioral model of health services utilization. D'Aunno and others have proposed that stronger linkages between care systems and clinicians (eg, case management [CM]), "colocated" services, more formal referral arrangements) can increase the likelihood of addictions treatment entry and completion.[24] Specialty substance dependence treatment reduces the risk for relapse. Medical and psychosocial services can help to prevent relapse and can help decrease barriers to substance dependence treatment, but patients face substantial systemic and personal barriers to receiving these services, as well as in accessing addiction specialty treatment (eg, insurance problems, personal disorganization, care systems focused on patients with one problem, privacy issues, bureaucracy, motivation).[26] Integrative linkages can align these services. Stronger linkage mechanisms increase the probability that patients will obtain needed services. The strongest linkage mechanism is on-site service delivery. To go beyond this structural solution, we conceptualize integrated, professionally delivered case management as a key component of CDM and as a health system-enabling resource as described in Andersen's model.[25] This resource helps the patient to increase recognition of needs (eg, substance dependence treatment, attention to medical and psychosocial needs) and to improve the use of health services ("effective access") that improve health status. Friedmann et al.[27,28] reported that drug use decreased among addiction treatment patients who had comprehensive services matched with identified needs. Thus, effective matching of service to need, one goal of CDM, can be a useful addiction treatment practice. In addition to these common elements that address systems issues and health and social problems, CDM tailored for addiction would include individual addiction-specific interventions. These elements are chosen based on their proven efficacy, their appropriateness for patients who have not entered treatment, their compatibility with health care organization theories, and primary care setting constraints. These addiction treatment components are each supported by their own theoretical frameworks: motivational interviewing by theories of behavior change[29,30]; case management by therapeutic alliance and enhancing receipt of needed services[31,32]; medication by theories of neuronal receptor involvement in alcohol and drug dependence[33,34]; complete health (medical, psychologic, social) and needs assessment, feedback, and advice delivered by an empathic clinician by the biopsychosocial model of addictions requiring pharmacotherapy, psychosocial support and services[35]; management of detoxification to avoid substance use to relieve withdrawal symptoms[36,37]; and relapse prevention in primary care.[38,39] Based on recent outcomes research, CDM for substance dependence could base recommendations for self-help involvement on an egalitarian model (offer to all regardless of individual need factors) and a need-based model for additional services.[40]

Finally, researchers have made a case for extended recovery monitoring interventions for alcohol and drug disorders.[41] In a randomized trial of patients with alcohol or cocaine dependence, 3 months of weekly telephone monitoring was added to group counseling sessions started after completion of intensive outpatient treatment. Total abstinence was more common 2 years later in the telephone monitoring group.[41] Similarly, other researchers have noted benefits from "recovery management check-ups."[42]

APPLICATION OF CDM TO ADDICTIONS

Wagner and colleagues proposed the chronic care model and the elements of effective chronic disease management (Table 1).[4–6,43,44] How should these elements work if applied to addiction care? As shown in Figure 1, in CDM, clinicians are expert, identify problems (disease of interest as well as medical, social, emotional), share information with patients and teach problem solving skills. Patients identify problems, set goals, and change behavior based on internal motivation. Multidisciplinary teams, including a nurse care manager, social worker and clinicians with expertise in the disease of interest, and expertise in common comorbidities, can spend time with the patient, coordinate with primary care physicians (PCPs), address necessary releases of information, and facilitate specialist referrals, provide access to community resources, implement evidence-based protocols, encourage self-management, and be proactive about follow-up. Information can be shared across team members, primary care clinicians, and specialists by using electronic records creating virtual colocation of care even when clinicians are in separate physical locations. This same information, when aggregated in a registry, can support the attention of the team to individual patients who have not received needed care and to clinical outcomes. *In sum, an informed, motivated patient and a prepared, proactive team and delivery system lead to optimal chronic disease care and improved outcomes.*

How could CDM elements be implemented in primary care for addiction? Figure 2 categorizes the specific elements of a proposed CDM intervention for substance (alcohol or drug) dependence into 3 areas: systems; medical, psychiatric and social problems; and addiction specific components and related outcomes. The systems changes follow directly from the elements in Table 1.

This approach is needed in part because of the current fragmentation of the treatment system, a system that only infrequently incorporates efficacious treatment elements. Although addiction treatment services often exist, and may be available in the sense that services are covered by insurance or grant funded, many barriers still prevent most patients from accessing these services. In fact, specialty services are not truly accessible at the time or in the settings in which patients are present. Chronic disease management, involving strong linkages within and between systems of care, integrated case management, and known effective addictions treatment components delivered under one roof, is hypothesized to increase the likelihood that patients will access and receive effective treatment for substance dependence problems, re-engage in care when they drop out, improve utilization of medical and other health services, and be less likely to suffer consequences of alcohol and drug use. These hypotheses are based on a careful review of the literature that we present in the section that follows regarding the chronic nature of addiction, fragmentation of care, suboptimal access to effective addiction care, and evidence for the individual components of the proposed CDM model.

ADDICTION: A CHRONIC DISEASE WITH PSYCHIATRIC AND MEDICAL COMORBIDITY

Whereas current care utilization is episodic (ie, detoxification only), the course of substance dependence often is chronic, particularly in people who seek and receive treatment. In a population-based sample, not a treatment sample, the mean duration of an episode of alcohol dependence was 3.7 years, and 28% of people had more than 1 episode (average, 5 episodes).[45] As with chronic medical illnesses, addiction is associated with chronic physiologic changes, a relapsing course, no "cure," variable adherence to care, and the need for ongoing care.[46,47] In addition to social, legal and family consequences, medical, and psychiatric disorders (eg, depression) are common (ie, 30–50%) and can be triggers for relapse.[38,48–54] When patients seek care for substance dependence, they are more likely than others to have myriad conditions: injury/overdose, anxiety, depression, psychosis, low back pain, headache, arthritis, asthma, acid-related peptic disorders, chronic obstructive pulmonary disease, hepatitis C, hypertension, alcoholic gastritis, diseases of the pancreas, and cirrhosis.[55,56] Because addiction can decrease medication adherence and other self-care, the care of these other medical and psychiatric conditions becomes more complex.[57]

FRAGMENTED ADDICTION TREATMENT

Elements of the proposed model of CDM provide strategies to reduce entry barriers to effective addiction treatment and to promote continuous patient engagement in some level of care. The current treatment system is fragmented among acute medical and specialty alcohol and drug services with little coordination.[58–60] Furthermore, few patients in recovery report having ever had formal treatment[61] and only 18% of adults with addiction report seeking mental health or addictions treatment annually.[19] Detoxification is a possible entry point to the treatment system but the missing linkage from detoxification, not a treatment itself, to treatment is recognized at the national level.[62] Privacy protection may interfere with coordinating linkage between treatment systems. However, the period after detoxification is a time of crisis during which mortality is substantial.[63] Barriers to treatment entry and continuous care include patient, treatment program, and systems factors. True accessibility[64] or "effective access"[25] is frequently missing, and only half of individuals entering treatment complete care episodes even

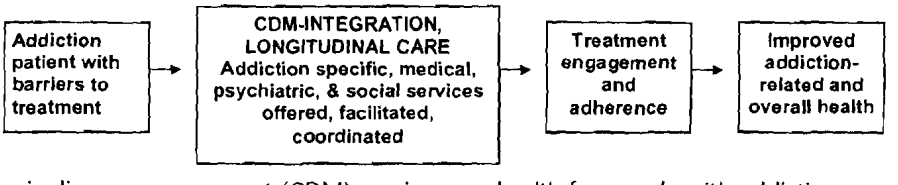


FIGURE 1. How chronic disease management (CDM) can improve health for people with addiction.

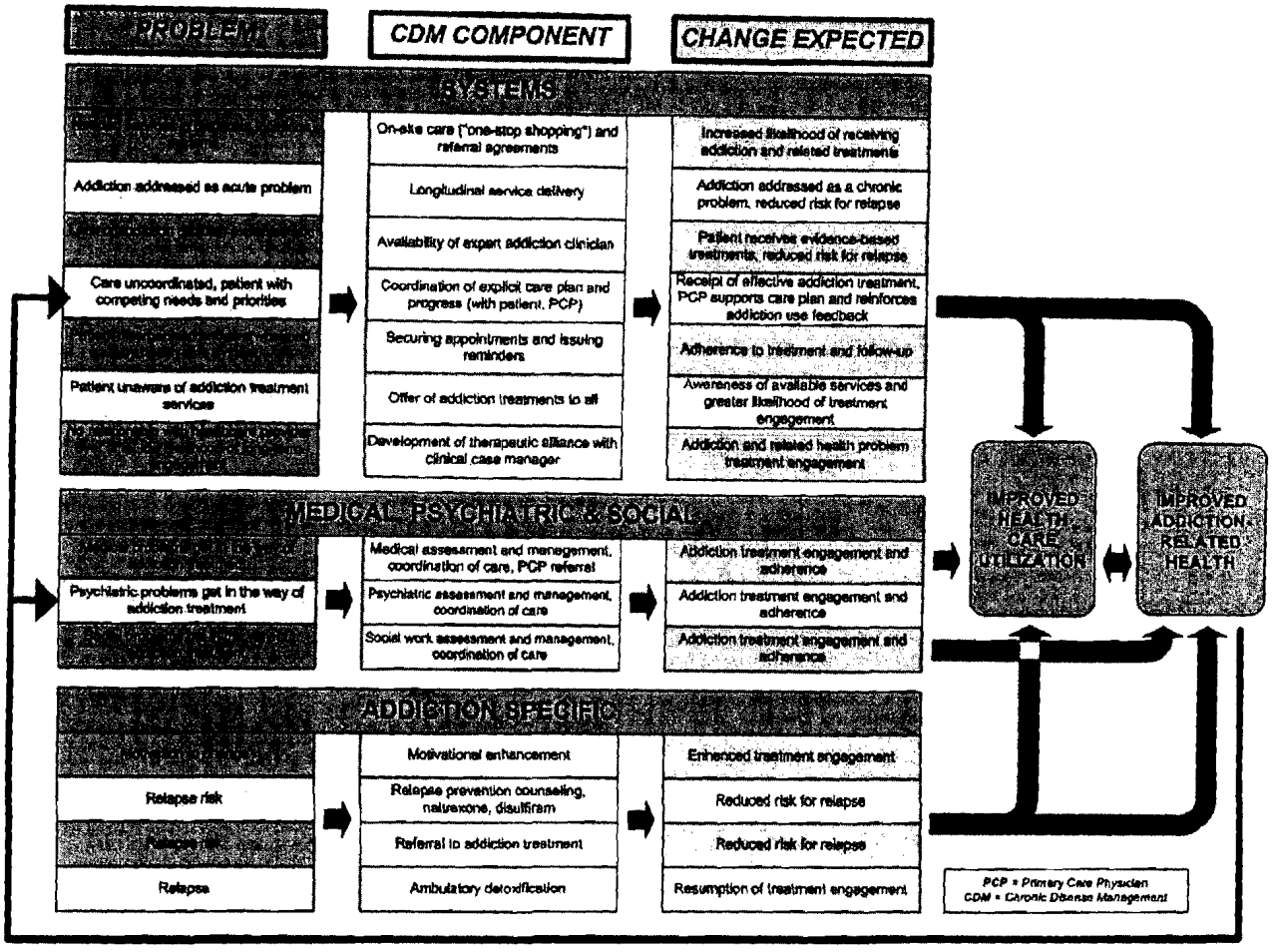


FIGURE 2. How chronic disease management (CDM) components address specific problems to lead to improved health.

though current programs are generally short-term.[65] Low rates occur despite the fact that financial barriers may be low, addiction treatment often is available without health insurance,[66] and persons with addictions generally live close to treatment programs, an average of 3.2 miles, although this availability is not consistent.[67] The treatment system may not be offering what the patient wants or can use. Patient factors include motivation, employment issues, attitudes, and beliefs,[68–71] and other comorbid medical and psychiatric problems can interfere with access to care. For example, a treatment program may preclude patients from concomitant use of psychiatric medication. These barriers can make "usual care," effectively no care, or, at best, suboptimal care for alcohol and drug dependence. Elements of CDM have the potential to address many of these barriers (Table 1; Figs. 1 and 2).

UNDERUTILIZED EFFICACIOUS TREATMENTS

Case management, pharmacotherapy (eg, naltrexone and buprenorphine), brief interventions, social skills training, community reinforcement approaches, behavior contracting, motivational interviewing, motivational incentives, and marital therapy are effective treatment elements that are currently underutilized for which use could be increased by CDM.[52,72–80] Pharmacological treatments are underused in the addiction system, which emerged outside of medical care settings. The orientation of the system is toward short-term interventions.[81] Treatment philosophy, particularly 12-step orientation, may run counter to any psychotropic drug use.[82,83] Health professionals have varied medical skills and qualifications (eg, ability to prescribe or administer medication[84]), and many lack both familiarity with addiction medications and ancillary support for its management.[85] Nevertheless, new interventions for addiction treatment particularly well suited to medical settings include "medical management" as tested in the COMBINE study[86] and used in the BRENDA (Biopsychosocial evaluation, Report, Empathy, Needs, Direct advice, Assessment) approach. BRENDA has primarily been used to support pharmacotherapy of alcoholism. Although not yet tested in a randomized trial nor used for patients with drug dependence, controlled studies have found use of BRENDA to be associated with improved medication adherence, retention in treatment, and improved clinical outcomes.[35,87,88] Preliminary data from primary care studies suggest that training to use such guidelines leads to clinician-based discussions about relapse risk.[38,89]

INSUFFICIENT LINKAGE OF ADDICTION TREATMENT AND PRIMARY MEDICAL CARE

Options for linkage of addiction and primary medical care treatment have primarily included distributive approaches, in which patients in one system are referred to another resulting in tenuous links.[59] Few patients with addictions access medical care and informal referrals from addictions treatment do not increase access.[24,90] Conversely, patients in medical settings are often not identified or referred to addiction treatment.[91–93] In 1991, a Federal conference concluded that better linkage should be pursued by colocating services and improving the effectiveness of referrals.[94] This conclusion remains relevant today.

A parallel problem exists for primary care providers accessing mental health services, an element of care also in short supply and often comorbid with substance dependence.[92,95] Primary care physicians have been challenged to play a larger role in addressing patients' mental health needs.[96–100] Some clinical models use a mental health team in the primary care clinic to rapidly evaluate and stabilize patients, and educate the primary care clinical staff.[99,100] In these integrated approaches receipt of mental health services increase, referrals decrease, patients benefit by continuing to be treated by their primary care physicians, and physicians benefit from additional support and training from mental health professionals. A quality improvement initiative for depression care management in primary care practices enhanced effectiveness and outcomes, leading an editorialist to conclude that "Evidence that depression outcomes can be improved through systematic changes in delivery of care is now compelling."[101,102] In another model, a Veterans Affairs mental health clinic[103] successfully integrated medical care emphasizing preventive measures, patient education, and close collaboration with mental health providers and patients had significant improvements in quality and outcomes of medical care. Mechanic[104] notes that although few systems successfully integrate care at the clinical level, simply having a clinician meet a community provider to whom patients with schizophrenia would be linked improved continuity of care and symptom improvement.[105] These mental health studies demonstrate the importance of discrete systemic innovations to improve access and clinical and utilization outcomes. Thus, although interorganizational integration and coordination of care has been difficult to achieve, when achieved, outcomes are improved.[106–108]

Intraorganizational interventions, such as CDM, have been proven feasible and effective and form the basis of the proposed model of CDM for addictions. Chronic disease management improves patient adherence to treatments and disease control compared with usual care and relies on patient education and reminders, and clinician education and feedback.[109–111] Compared with usual care, CDM interventions focus on a disease but attend to comorbidity. A major strength and promise of such interventions is improvement in the disease of focus as well as comorbidities.[101–103,112–115]

Although no trials of CDM for addiction appear in the literature, systematic reviews have identified numerous controlled studies of CDM for psychiatric illnesses and inform the choice of effective elements for addictions. For example, a large randomized trial of a depression CDM intervention found that CDM patients were more likely to receive depression treatments and have a significant reduction in depressive symptoms and functional impairment[116] and improved arthritis pain and functional outcomes.[117] In this CDM intervention, patients received a 20-minute educational video and booklet about depression, a visit with a trained depression nurse or psychologist care manager in primary care, underwent medical and psychosocial assessment, and were asked about their treatment preferences. The team suggested antidepressant medication (80% had 1 trial of medication) or brief psychotherapy adapted for the primary care setting delivered by the case manager (30% received this counseling). Thus, the team offered services tailored to patients' needs and preferences rather than attempting to provide a specific treatment to all subjects regardless of preference or need. Patients were contacted by telephone (mean, 6 times) or visited the clinic every other week (mean, 9 visits). When symptoms persisted, the team, including a specialist (psychiatrist), intervened. The depression CDM intervention included focus on a medical condition if the patient wished.

Rost et al.[118] randomized 211 adults with major depression to usual primary medical care or to a CDM intervention group. An "initial intervention" was a visit with a trained office nurse to assess depressive symptoms, provide education about treatment options, address readiness to engage in treatment, and arrange follow-up. In the "continuing intervention," nurse care managers phoned patients to encourage adherence and to advise them to raise problems with their physicians. The content was tailored to whether the patient was symptomatic. Nurse care managers provided physicians with reports of patients' symptoms and treatments, and a psychiatrist provided treatment guidelines. Chronic disease management increased remission of depression and improved emotional and physical functioning.

Simon et al.[119] compared CDM with usual medical care for depression. Chronic disease management included telephone assessments of depressive symptoms and treatments, calls to those who discontinued treatments, assistance with appointments, reports to primary care physicians on patient progress with depression treatment recommendations, and suggestions to contact patients to schedule follow-up visits as needed. A psychiatrist supervised the CDM team. Chronic disease management improved prescription of appropriate doses of medication and decreased depressive symptoms and the incidence of major depression compared with usual care.

Hedrick et al.[120] tested CDM for depression compared with usual primary care, in which psychiatrists were available for referrals in primary care. Care was by a multidisciplinary team in primary care consisting of a psychologist, psychiatrist, social workers, and a psychology technician. The team implemented evidence-based care (antidepressants and 6 sessions of cognitive behavioral therapy), communicated with primary care physicians, took patient preferences into account, proactively monitored patient symptoms and treatment barriers, and provided patient education. Chronic disease management increased the proportion of patients receiving effective therapies for depression and was associated with

improved depressive symptoms and mental health-related quality of life.

Because substance dependence is associated with significant medical illness and cost consequences, the evidence regarding CDM for medical illnesses also is relevant. Chronic disease management for heart failure, diabetes, arthritis and asthma, in randomized, controlled studies[121-126] and in systematic reviews of more than 100 trials[43,44,127] leads to clinical and functional improvement, decreased hospitalizations, treatment adherence, and patient satisfaction. Health systems adopting these programs report improved outcomes.[44] Of note, Whellan et al.'s[126] intervention, in addition to addressing the target chronic disease of heart failure directly, also included "other strategies targeted at optimizing the control of concomitant illnesses that may worsen" heart failure. A lesson for addictions care might be to include care for common comorbid medical and psychiatric illnesses in substance dependence CDM.

Evidence for effectiveness of CDM for psychiatric and medical illnesses is strong. Because addiction has similarities with these chronic illnesses, CDM has potential for improving addiction outcomes.

ELEMENTS OF CDM FOR ADDICTION PROVEN EFFECTIVE

A Medline search from 1966 through late 2007 for "disease management" and "alcoholism" or "drug dependence" yielded few results, none of which included controlled trial evidence for the effectiveness of CDM for addiction. Given the lack of relevant published studies, we reviewed the evidence for 2 key components, which when combined would reflect on the potential benefit of CDM when studied in patients with addiction: a) integrated case management delivered by professionals, and b) integrative linkage of medical, psychosocial, and alcohol and drug dependence care.

Case management coordinates and links patients with appropriate services to address specific needs across systems of care,[32,128,129] and in this model is delivered by a skilled nurse or social worker. Case management includes patient assessment, care planning and coordination, linkage to services, outcome monitoring, and advocacy for patients, in a single point of contact, for addiction, medical, and other services (eg, family services, self-help groups, insurance, food, housing, transportation, and employment). In alcohol and drug treatment practice, the CM definition is quite varied,[32,130,131] may be delivered by paraprofessionals or peers, and not all models are effective. When CM and addiction treatment are delivered by one clinician, the approach is more effective than case management alone.[32] One key ingredient is therapeutic alliance.[132] This alliance impacts homelessness,[31] treatment participation, drinking,[133] and drug treatment retention and outcome,[134,135] particularly for those with more severe psychiatric problems.[136]

Case management can increase linkage from substance abuse treatment to primary medical care.[48] Case management also can decrease relapse and increase retention in addictions treatment.[137] McLellan et al. compared outpatient group counseling twice per week with counseling and case management in a quasi-experimental study in patients admitted to addiction treatment.[15-20] Case management was associated with greater receipt of alcohol, medical, psychiatric, employment, and family services, and with less alcohol intoxication (and lower severity), and fewer days of psychiatric and medical problems.[75] In a similar study,[52] Case management improved alcohol use, medical, employment, legal, and family status. Stout et al.[130] compared case monitoring aftercare delivered by skilled clinicians (case management plus reassessment and ongoing advice) for patients with alcohol abuse or dependence discharged from day hospital to standard referrals to outpatient follow-up. Case monitors met with subjects for 30 minutes and then by telephone monthly or less for 2 years. The interactions included constant reassessment, were supportive and nonjudgmental, and addressed substance use and other major life problems by referral. Recommendations depended on patient needs. Preliminary results were a 50% decrease in heavy drinking and fewer emergency visits in the CM group.[41,138] Dennis et al.[42] found that quarterly case management delivered by phone for patients in early recovery led to more appropriate treatment utilization.

A second element involves organizational restructuring: *integration and "colocation" of services* to achieve integration and continuity of care.[139-142] Studies of colocation have found that patients with addictions who receive both regular addiction and medical care were less likely to be hospitalized than those who received one or neither service,[143] and on-site medical service provision, transportation, and CM increased receipt of medical services.[144-150] Friedmann et al.[151] found that provision of primary medical care by off-site referral or on-site at drug treatment programs, compared with no such mechanism, reduced emergency, and hospital utilization.

Furthermore, on-site primary care at addiction programs has been associated with reduced addiction severity.[152] In a randomized trial, patients receiving on-site medical, psychiatric, employment, and family services had less opiate use, and improved medical, employment, legal and psychiatric outcomes.[74] Women with psychiatric problems were more likely to complete outpatient addictions care when offered psychiatric care.[153] In a trial that randomized veterans without primary care who were entering substance abuse treatment with a chronic medical condition to receive primary medical care either on-site or off-site, on-site care increased access to primary care and addiction treatment retention.[154]

On-site alcohol and drug treatment in primary care also can improve alcohol and drug use outcomes and be safe and effective.[155-159] In clinical trials, naltrexone for alcoholism was efficacious when given with primary care management.[160] Similar findings of success of acamprosate in primary care have been reported.[161] Furthermore, one study reported that 78% of patients receiving office-based buprenorphine for opiate dependence remained in care compared with 52% of patients in a traditional drug treatment center.[158]

In a unique model, Weisner et al.[162] randomized 592 adults to usual, separate primary care, or *integrated primary care at an addictions treatment program* by 3 primary care physicians with specialty addictions training, a medical assistant, and 2 nurses. There were no overall differences in

abstinence, but in a subgroup of patients with substance abuse-related medical conditions (57%), on-site care was associated with increased abstinence at 6 months. In a randomized trial in a special alcohol clinic for veterans,[163] the integrated care group was more likely to be abstinent than a usual care group (74% versus 49% 30-day abstinence).[164] The study intervention was focused on alcoholism but included substantial attention to comorbidity. The intervention was an initial thorough inpatient evaluation by a multidisciplinary team who developed a care plan to reduce alcohol severity and remission of related medical conditions. The plan included monthly primary care visits to review drinking and medical problems at a frequency indicated by clinical status and feedback of blood test results to encourage abstinence. Mental health and social services and more intensive alcohol treatment were provided on-site when necessary. Patients were contacted when they missed appointments. In a recent prospective study,[165] patients with alcohol dependence were referred to Alcoholics Anonymous; when most refused to attend, the study provided monthly extensive visits with a medical nurse who was available for telephone consults, and brief visits with a gastroenterologist. Drinking decreased from 16 to 2.5 drinks per day.

Andersen et al.[166] studied 45 adults cared for by a nurse care manager who addressed both their HIV and substance abuse, accompanying patients to physician visits and facilitating integration of medical and substance abuse treatment recommendations. Addiction severity and health-related quality of life improved significantly in this sample during 6 months. Bartels et al.[167] studied 2022 elderly patients with a mental health disorder and/or at-risk drinking, randomizing them to integration and colocation of mental health and substance abuse services in primary care, or to facilitated referral, including scheduling and payment, and transportation, to specialty mental health, or substance abuse clinics. The integrated model was associated with greater attendance at mental health and substance abuse treatment. This body of research, both randomized controlled trials and cohort studies, supports the concept that integration of addiction, mental health, and medical services yields improvements in adherence to care, severity of substance use, and appropriate utilization of services.

CONCLUSIONS

Substance dependence is a common and costly chronic illness associated with medical and psychiatric comorbidity. Treatment can be efficacious when it is actually received by patients. But the current system of care is fragmented, not coordinated, and does not always include proven efficacious treatments. Patient motivation and coexisting health and social problems are barriers to receipt of effective treatment. Integrated and coordinated care, which simultaneously addresses patient motivation and needs across health domains, provides efficacious addiction treatments and facilitates effective access to other treatment. This integrated care may increase the likelihood that care is received and that addiction-related and other clinical outcomes improve. The World Health Organization called the management of chronic conditions "one of the greatest challenges facing health care systems throughout the world" and recommended building integrated health care as an essential part of the solution.[168] Chronic disease management is a relatively new model to care for chronic psychiatric and medical illnesses and has not been fully applied or disseminated for alcohol or drug dependence. In fact, the leading and latest literature on the topic, prompted by calls from the Institute of Medicine, is silent regarding addictions.[169] More recently, the Institute of Medicine has again, and more specifically, called for improvements in the quality of care for substance use conditions.[170] Chronic disease management is one way to advance this agenda.

Chronic disease management shows promise as an effective strategy for managing substance dependence. It is critical to test the effectiveness of CDM integrated in a primary care setting for substance dependent patients, because this approach can take advantage of the fact that many patients with addictions attend primary care yet do not receive specialty care for their addictions. The current fragmented health service delivery models are limited in many ways for patients with the chronic illness of substance dependence. While we await studies of the effectiveness of CDM in primary care, elements of CDM could be implemented now.

In 1996, we judged from a review of the literature that linking people with addictions with primary medical care[59] held promise, and later we detailed the potential benefits.[58] In a randomized trial, we demonstrated that multidisciplinary assessment and referral increased linkage of people with addictions to primary medical care but found that simple linkage was not enough to improve health.[171] Based on review of the latest literature, the evidence suggests that services delivery models that include case management and integrated care and are modeled on chronic disease (care) management hold promise for improving the care received by people with substance dependence.

ACKNOWLEDGMENTS

The authors thank the Health Evaluation and Linkage to Primary care (HELP) and the Addiction Health Evaluation and Disease management (AHEAD) study teams for their assistance in thinking broadly about the issues covered in this paper.

REFERENCES

1. Humphreys K, Tucker JA. Toward more responsive and effective intervention systems for alcohol-related problems. *Addiction.* 2002;97: 126–132.
2. Trask PC, Schwartz SM, Deaner SL, et al. Behavioral medicine: the challenge of integrating psychological and behavioral approaches into primary care. *Eff Clin Pract.* 2002;5:75–83.
3. Weisner C, Schmidt LA. Expanding the frame of health services research in the drug abuse field. *Health Serv Res.* 1995;30:707–726.
4. Von Korff M, Gruman J, Schaefer J, et al. Collaborative management of chronic illness. *Ann Intern Med.* 1997;127:1097–1102.
5. Wagner EH. The role of patient care teams in chronic disease management. *BMJ.* 2000;320:569–572.
6. Wagner EH, Austin BT, Von Korff M. Organizing care for patients with chronic illness. *Milbank Q.* 1996;74:511–544.

7. Disease Management Association of America. Available at: http://www.dmaa.org/dm_definition.asp. Accessed July 14, 2007.
8. Institute of Medicine. Crossing the Quality Chasm: A New Health System for the 21st Century. Committee on Quality of Health Care in America. Washington, DC: National Academy Press; 2001.
9. Hunter DJ, Fairfield G. Disease management. *BMJ.* 1997;315:50–53.
10. Department of Health Services. Medicare program; voluntary chronic care improvement under traditional fee-for-service Medicare. Federal Register 2004;69:22065–22079.
11. National Committee for Quality Assurance. Disease management standards and guidelines for accreditation. Draft standards for public comment. Available at: http://www.NCQA.org. Accessed July 17, 2007.
12. Bodenheimer T. Disease management: promises and pitfalls. *N Engl J Med.* 1999;340:1202–1205.
13. Bodenheimer T. Disease management in the American market. *BMJ.* 2000;320:563–566.
14. DeBusk RF, West JA, Miller NH, et al. Chronic disease management: treating the patient with disease(s) vs treating disease(s) in the patient. *Arch Intern Med.* 1999;159:2739–2742.
15. Kendrick T. Depression management clinics in general practice? Some aspects lend themselves to the mini-clinic approach. *BMJ.* 2000;320:527–528.
16. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 4th ed. Washington, DC: American Psychiatric Association; 1994.
17. Substance Abuse and Mental Health Services Administration. Office of Applied Studies. Results from the 2005 National Survey on Drug Use and Health: National Findings (Office of Applied Studies, NSDUH Series H-30, DHHS Publication No. SMA 06-4194). Rockville, MD. Available at: http://oas.samhsa.gov/NSDUH/2k5NSDUH/2k5results.htm#7. 3 Accessed July 17, 2007.
18. Cartwright WS. Economic costs of drug abuse in the United States. *J Mental Health Policy Econ.* 1999;2:133–134.
19. Green-Hennessey S. Factors associated with receipt of behavioral health services among persons with substance dependence. *Psychiatr Serv.* 2002;53:1592–15798.
20. Harwood HJ, Fountain D, Livermore G. The economic costs of alcohol and drug abuse in the United States-1992. Washington, DC: National Institute on Drug Abuse, National Institute on Alcohol Abuse and Alcoholism; 1998.
21. Harwood HJ. Updating estimates of the economic costs of alcohol abuse in the United States. Washington, DC: National Institute on Alcohol Abuse and Alcoholism; 2000.
22. National Institute on Alcohol Abuse and Alcoholism. Drinking in the United States: main findings from the 1992 National Longitudinal Alcohol Epidemiologic Survey (NLAES). Bethesda, MD: U.S. Department of Health and Human Services; 1998.
23. Grant BF, Dawson DA, Stinson FS, et al. The 12-month prevalence and trends in DSM-IV alcohol abuse and dependence: United States, 1991–1992 and 2001–2002. *Drug Alcohol Depend.* 2004;74:223–234.
24. D'Aunno T. Linking substance-abuse treatment and primary health Care. In: V Egertson JA, Fox DM, Leshner AI, eds. Treating Drug Abusers Effectively. Cambridge, MA: Blackwell; 1997:311–331.
25. Andersen RM. Revisiting the behavioral model and access to medical care: does it matter? *J Health Soc Behav.* 1995;36:1–10.
26. Teitelbaum MA, Walker M, Gabay M, et al. Barriers to the delivery of integrated primary care services and substance abuse treatment: case studies of nine linkage program projects. Washington, DC: Health Resources and Services Administration and ABT Associates, Inc.; 1992.
27. Friedmann PD, Hendrickson JC, Gerstein D, et al. The effect of matching comprehensive services to patients' needs on drug use improvement in addiction treatment. *Addiction.* 2004;99:962–972.
28. Saitz R, Horton NJ, Larson MJ, et al. Primary medical care and reductions in addiction severity: a prospective cohort study. *Addiction.* 2005;100:70–78.
29. Miller WR, Rollnick S. Motivational Interviewing. Preparing People to Change Addictive Behavior. New York: The Guilford Press; 1991.
30. Prochaska JO, DiClemente CC, Norcross JC. In search of how people change. Applications to addictive behaviors. *Am Psychol.* 1992;47:1102–1114.
31. Chinman MJ, Rosenheck R, Lam JA. The case management relationship and outcomes of homeless persons with serious mental illness. *Psychiatr Serv.* 2000;51:1142–1147.
32. Ridgely SM, Willenbring ML. Application of case management to drug abuse treatment: overview of models and research issues. In: Ashery R, ed. Progress and Issues in Case Management. Rockville, MD: National Institute on Drug Abuse; 1992:12–33.
33. Litten RZ, Allen JP. Advances in development of medications for alcoholism treatment. *Psychopharmacology (Berl).* 1998;139:20–33.
34. Dole VP. Implications of methadone maintenance for theories of narcotic addiction. *JAMA.* 1988;260:3025–3029.
35. Volpicelli JR, Pettinati HM, McLellan AT, et al. Combining Medication and Psychosocial Treatments for Addictions. The BRENDA Approach. New York: The Guilford Press; 2001.
36. Edwards G. Withdrawal symptoms and alcohol dependence: fruitful mysteries. *Br J Addict.* 1990;85:447–461.
37. Kosten TR, O'Connor PG. Management of drug and alcohol withdrawal. *N Engl J Med.* 2003;348:1786–1795.
38. Friedmann PD, Saitz R, Samet JH. Management of adults recovering from alcohol or other drug problems: relapse prevention in primary care. *JAMA.* 1998;279:1227–1231.
39. Friedmann PD, Rose J, Hayaki J, et al. Training primary care clinicians in maintenance care for moderated alcohol use. *J Gen Intern Med.* 2006;21:1269–1275.
40. Moos RH, Moos BS. The interplay between help-seeking and alcohol-related outcomes: divergent processes for professional treatment and self-help groups. *Drug Alcohol Depend.* 2004;75:155–164.
41. McKay JR. Is there a case for extended interventions for alcohol and drug use disorders? *Addiction* 2005;100:1594–1610.
42. Dennis M, Scott CK, Funk R. An experimental evaluation of recovery management checkups (RMC) for people with chronic substance use disorders. *Eval Program Planning.* 2003;26:339–352.
43. Bodenheimer T, Lorig K, Holman H, et al. Patient self-management of chronic disease in primary care. *JAMA.* 2002;288:2469–2475.
44. Bodenheimer T, Wagner EH, Grumbach K. Improving primary care for patients with chronic illness. *JAMA.* 2002;288:1775–1779.
45. Hasin DS, Stinson FS, Ogburn E, et al. Prevalence, correlates, disability, and comorbidity of DSM-IV alcohol abuse and dependence in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Arch Gen Psychiatry.* 2007;64:830–842.
46. McLellan AT. Have we evaluated addiction treatment correctly? Implications from a chronic care perspective. *Addiction.* 2002;97:249–252.
47. 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.
48. Brindis C, Pfeffer R, Wolfe A. A case management program for chemically dependent clients with multiple needs. *J Case Manag.* 1995;4:22–28.
49. Brooner RK, King VL, Kidorf M, et al. Psychiatric and substance use comorbidity among treatment-seeking opioid abusers. *Arch Gen Psychiatry.* 1997;54:71–80.
50. Hasin D, Liu X, Nunes E, et al. Effects of major depression on remission and relapse of substance dependence. *Arch Gen Psychiatry.* 2002;59:375–380.
51. Kessler RC, 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.
52. McLellan AT, Hagan TA, Levine M, et al. Does clinical case management improve outpatient addiction treatment. *Drug Alcohol Depend.* 1999;55:91–103.
53. Regier DA, Farmer ME, Rae DS, et al. Comorbidity of mental disorders with alcohol and other drug abuse. Results from the Epidemiologic Catchment Area (ECA) Study. *JAMA.* 1990;264:2511–2518.
54. Ziedonis D, Brady K. Dual diagnosis in primary care. Detecting and treating both the addiction and mental illness. *Med Clin North Am.* 1997;81:1017–1036.
55. Mertens JR, Lu YW, Parthasarathy S, et al. Medical and psychiatric conditions of alcohol and drug treatment patients in an HMO: comparison with matched controls. *Arch Intern Med.* 2003;163:2511–2517.

56. Compton WM, Thomas YF, Stinson FS, et al. Prevalence, correlates, disability, and comorbidity of DSM-IV drug abuse and dependence in the United States: results from the national epidemiologic survey on alcohol and related conditions. *Arch Gen Psychiatry.* 2007;64:566–576.
57. Golin CE, Liu H, Hays RD, et al. A prospective study of predictors of adherence to combination antiretroviral medication. *J Gen Intern Med.* 2002;17:756–765.
58. 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;161:85–91.
59. Samet JH, Saitz R, Larson MJ. A case for enhanced linkage of substance abusers to primary medical care. *Subst Abuse.* 1996;17:181–192.
60. Weisner C, McCarty D, Schmidt L. New directions in alcohol and drug treatment under managed care. *Am J Managed Care.* 1999;5:SP57–SP69.
61. Brown RL, Saunders LA, Bobula JA, et al. Remission of alcohol disorders in primary care patients. Does diagnosis matter? *J Fam Pract.* 2000;49:522–528.
62. Mark T, Dilionardo J, Chalk M, Coffey R. Substance abuse detoxification: improvements needed in linkage to treatment (SMA-02-3728). Rockville, MD: Substance Abuse and Mental Health Services Administration; 2002.
63. Saitz R, Gaeta J, Cheng DM, et al. Risk of mortality during four years after substance detoxification in urban adults. *J Urban Health.* 2007; 84:272–282.
64. Institute of Medicine. Broadening the Base of Treatment for Alcohol Problems. Washington, DC: National Academy Press; 1990.
65. Substance Abuse and Mental Health Services Administration. Office of Applied Studies. Treatment completion in the Treatment Episode Data Set (TEDS), The Drug and Alcohol Services Information System (DASIS), Rockville, MD: January 30, 2003. Available at: http://www.oas.samhsa.gov/2k3/TXcompleters/TXcompleters.cfm. Accessed July 17, 2007.
66. Substance Abuse and Mental Health Services Administration. Health insurance status of admissions for substance abuse treatment: 1999. Available at: http://www.oas.samhsa.gov/2k2/insuranceTX/insuranceTX.cfm. Accessed July 17, 2007.
67. Substance Abuse and Mental Health Services Administration. Office of Applied Studies. The National Household Survey on Drug Abuse (NHSDA), Drug and Alcohol Services Information System (DASIS). Distance to substance abuse treatment facilities among those with alcohol dependence or abuse. Available at: http://www.oas.samhsa.gov/2k2/distance/distance.htm. Accessed July 17, 2007.
68. Grant BF. Barriers to alcoholism treatment: reasons for not seeking treatment in a general population sample. *J Stud Alcohol.* 1997;58:365–371.
69. Joe GW, Simpson DD, Broome KM. Effects of readiness for drug abuse treatment on client retention and assessment of process. *Addiction.* 1998;93:1177–1190.
70. Joe GW, Simpson DD, Greener JM, et al. Integrative modeling of client engagement and outcomes during the first 6 months of methadone treatment. *Addict Behav.* 1999;24:649–659.
71. Weisner C, Mertens J, Tam T, et al. Factors affecting the initiation of substance abuse treatment in managed care. *Addiction.* 2001;96:705–716.
72. Miller WR, Wilbourne PL. Mesa Grande: a methodological analysis of clinical trials of treatments for alcohol use disorders. *Addiction.* 2002; 97:265–277.
73. Johnson RE, Chutuape MA, Strain EC, et al. A comparison of levomethadyl acetate, buprenorphine, and methadone for opioid dependence. *N Engl J Med.* 2000;343:1290–1297.
74. McLellan AT, Arndt IO, Metzger DS, et al. The effects of psychosocial services in substance abuse treatment. *JAMA.* 1993;269:1953–1959.
75. McLellan AT, Hagan TA, Levine M, et al. Supplemental social services improve outcomes in public addiction treatment. *Addiction.* 1998;93:1489–1499.
76. Higgins ST, Petry NM. Contingency management. Incentives for sobriety. *Alcohol Res Health.* 1999;23:122–127.
77. Rawson R. A. Treatment for stimulant use disorders, Treatment Improvement Protocol (TIP) Series 33. Available at: http://ncadi.samhsa.gov/govpubs/BKD289/33f.aspx. Accessed July 17, 2007.
78. Burke BL, Arkowitz H, Dunn C. The efficacy of motivational interviewing and its adaptations. What we know so far. In: Miller W, Rollnick S, eds. In: Motivational Interviewing. Preparing People for Change. 2nd ed. New York: Guilford Press; 2002:217–250.
79. Zweben A, Zuckoff A. Motivational interviewing and treatment adherence. In: Miller W, Rollnick S, eds. Motivational Interviewing. Preparing People for Change. 2nd ed. New York: Guilford Press; 2002: 299–319.
80. Fiellin DA, O'Connor PG. Clinical practice. Office-based treatment of opioid-dependent patients. *N Engl J Med.* 2002;347:817–823.
81. Thomas C, Wallack S, Swift D, et al. Barriers to adoption of naltrexone in alcoholism treatment: a three-state study. *Alcohol Clin Exp Res.* 2000;24:38A.
82. Knudsen HK, Ducharme LJ, Roman PM, et al. Buprenorphine diffusion: the attitudes of substance abuse treatment counselors. *J Subst Abuse Treat.* 2005;29:95–106.
83. Thomas CP, Wallack SS, Lee S, et al. Research to practice: adoption of naltrexone in alcoholism treatment. *J Subst Abuse Treat.* 2003;24:1–11.
84. Garbutt JC, Kranzler HR, O'Malley SS, et al. Efficacy and tolerability of long-acting injectable naltrexone for alcohol dependence: a randomized controlled trial. *JAMA.* 2005;293:1617–1625.
85. Ward LD, Shea JA, Laine C, et al. Willingness to provide primary care office-based methadone maintenance (PCOM): a qualitative study of clinic directors [abstract]. *J Gen Intern Med.* 2003;18:277.
86. Miller WR, Moyers TB, Arciniega L, et al. Training, supervision and quality monitoring of the COMBINE Study behavioral interventions. *J Stud Alcohol Suppl.* 2005;188–195.
87. Pettinati HM, Volpicelli JR, Pierce JD Jr, et al. Improving naltrexone response: an intervention for medical practitioners to enhance medication compliance in alcohol dependent patients. *J Addict Dis.* 2000;19: 71–83.
88. Kaempf G, O'Donnell C, Oslin DW. The BRENDA model: a psychosocial addiction model to identify and treat alcohol disorders in elders. *Geriatr Nurs.* 1999;20:302–304.
89. Friedmann PD, Ramsey S, Herman D, et al. *Maintenance care for alcohol problems in remission in primary care.* Presented at the 2002 National Meeting of the Association for Medical Education and Research on Substance Abuse. 2002, Washington, DC.
90. Laine C, Newschaffer C, Zhang D, et al. Models of care in New York State Medicaid substance abuse clinics. Range of services and linkages to medical care. *J Subst Abuse.* 2000;12:271–285.
91. Buchsbaum DG, Buchanan RG, Poses RM, et al. Physician detection of drinking problems in patients attending a general medicine practice. *J Gen Intern Med.* 1992;7:517–521.
92. Saitz R, Mulvey KP, Plough A, et al. Physician unawareness of serious substance abuse. *Am J Drug Alcohol Abuse.* 1997;23:343–354.
93. Saitz R, Mulvey KP, Samet JH. The substance abusing patient and primary care: linkage via the addiction treatment system? *Substance Abuse.* 1997;18:187–195.
94. Kunnes R, Niven R, Gustafson T, et al. Financing and payment reform for primary health care and substance abuse treatment. *J Addict Dis.* 1993;12:23–42.
95. Trude S, Stoddard JJ. Referral gridlock: primary care physicians and mental health services. *J Gen Intern Med.* 2003;18:442–449.
96. Arboleda-Florez J, Saraceno B. Mental health and primary care. *Can Med Assoc J.* 2001;164:1013–1014.
97. Dodds S, Blaney NT, Nuehring EM, et al. Integrating mental health services into primary care for HIV-infected pregnant and non-pregnant women: whole life- a theoretically derived model for clinical care and outcomes assessment. *Gen Hosp Psychiatry.* 2000;22:251–260.
98. Feingold A, Slammon WR. A model integrating mental health and primary care services for families with HIV. *Gen Hosp Psychiatry.* 1993;15:290–300.
99. Felker BL, Barnes RF, Greenberg DM, et al. Preliminary outcomes from an integrated mental health primary care team. *Psychiatr Serv.* 2004;55:442–444.
100. Kates N, Craven M, Crustolo AM, et al. Integrating mental health services within primary care. *Gen Hosp Psychiatry.* 1997;19:324–332.
101. Von Korff M. Can care management enhance integration of primary and specialty care? *BMJ.* 2004;329:605.

102. Dietrich A, Oxman TE, Williams JW, et al. Re-engineering systems for the treatment of depression in primary care: cluster randomised controlled trial. *BMJ.* 2004;329:602–605.
103. Druss BG, Rohrbaugh RM, Levinson CM, et al. Integrated medical care for patients with serious psychiatric illness: a randomized trial. *Arch Gen Psychiatry.* 2001;58:861–868.
104. Mechanic D. Policy challenges in improving mental health services: some lessons from the past. *Psychiatr Serv.* 2003;54:1227–1232.
105. Olfson M, Mechanic D, Boyer CA, et al. Linking inpatients with schizophrenia to outpatient care. *Psychiatr Serv.* 1998;49:911–917.
106. Glisson C, Hemmelgarn A. The effects of organizational climate and interorganizational coordination on the quality and outcomes of children's service systems. *Child Abuse Negl.* 1998;22:401–421.
107. Goldman HH, Morrissey JP, Ridgely MS. Evaluating the Robert Wood Johnson Foundation program on chronic mental illness. *Milbank Q.* 1994;72:37–47.
108. Bickman L. Implications of a children's mental health managed care demonstration evaluation. *J Ment Health Adm.* 1996;23:107–117.
109. Mueser KT, Corrigan PW, Hilton DW, et al. Illness management and recovery: a review of the research. *Psychiatr Serv.* 2002;53:1272–1284.
110. Weingarten SR, Henning JM, Badamgarav E, et al. Interventions used in disease management programmes for patients with chronic illness-which ones work? Meta-analysis of published reports. *BMJ.* 2002;325:925.
111. Badamgarav E, Weingarten SR, Henning JM, et al. Effectiveness of disease management programs in depression: a systematic review. *Am J Psychiatry.* 2003;160:2080–2090.
112. Katon W, Russo J, Von Korff M, et al. Long-term effects of a collaborative care intervention in persistently depressed primary care patients. *J Gen Intern Med.* 2002;17:741–748.
113. Katon WJ, Roy-Byrne P, Russo J, et al. Cost-effectiveness and cost offset of a collaborative care intervention for primary care patients with panic disorder. *Arch Gen Psychiatry.* 2002;59:1098–1104.
114. Roy-Byrne PP, Katon W, Cowley DS, et al. A randomized effectiveness trial of collaborative care for patients with panic disorder in primary care. *Arch Gen Psychiatry.* 2001;58:869–876.
115. Schoenbaum M, Unutzer J, Sherbourne C, et al. Cost-effectiveness of practice-initiated quality improvement for depression: results of a randomized controlled trial. *JAMA.* 2001;286:1325–1330.
116. Unutzer J, Katon W, Callahan CM, et al. Collaborative care management of late-life depression in the primary care setting: a randomized controlled trial. *JAMA.* 2002;288:2836–2845.
117. Lin EH, Katon W, Von Korff M, et al. Effect of improving depression care on pain and functional outcomes among older adults with arthritis: a randomized controlled trial. *JAMA.* 2003;290:2428–2429.
118. Rost K, Nutting P, Smith JL, et al. Managing depression as a chronic disease: a randomised trial of ongoing treatment in primary care. *BMJ.* 2002;325:934.
119. Simon GE, VonKorff M, Rutter C, et al. Randomised trial of monitoring, feedback, and management of care by telephone to improve treatment of depression in primary care. *BMJ.* 2000;320:550–554.
120. Hedrick SC, Chaney EF, Felker B, et al. Effectiveness of collaborative care depression treatment in Veterans' Affairs primary care. *J Gen Intern Med.* 2003;18:9–16.
121. Blonde L. Disease management approaches to type 2 diabetes. *Manag Care.* 2000;9:18–23.
122. McAlister FA, Lawson FM, Teo KK, et al. A systematic review of randomized trials of disease management programs in heart failure. *Am J Med.* 2001;110:378–384.
123. Rich MW, Beckham V, Wittenberg C, et al. A multidisciplinary intervention to prevent the readmission of elderly patients with congestive heart failure. *N Engl J Med.* 1995;333:1190–1195.
124. Riegel B, Carlson B, Kopp Z, et al. Effect of a standardized nurse case-management telephone intervention on resource use in patients with chronic heart failure. *Arch Intern Med.* 2002;162:705–712.
125. Vrijhoef HJ, Diederiks JP, Spreeuwenberg C. Effects on quality of care for patients with NIDDM or COPD when the specialised nurse has a central role: a literature review. *Patient Educ Couns.* 2000;41:243–250.
126. Whellan DJ, Gaulden L, Gattis WA, et al. The benefit of implementing a heart failure disease management program. *Arch Intern Med.* 2001;161:2223–2228.
127. Bodenheimer T, Wagner EH, Grumbach K. Improving primary care for patients with chronic illness: the chronic care model, Part 2. *JAMA.* 2002;288:1909–1914.
128. National Institute on Drug Abuse. Progress and Issues in Case Management. Rockville, MD: National Institute on Drug Abuse; 1992.
129. Siegal H. A. Comprehensive case management for substance abuse treatment. Rockville, MD: U.S. Department of Health and Human Services; 2001.
130. Stout RL, Rubin A, Zwick W, et al. Optimizing the cost-effectiveness of alcohol treatment: a rationale for extended case monitoring. *Addict Behav.* 1999;24:17–35.
131. Willenbring ML, Ridgely MS, Stinchfield R, et al. Application of case management in alcohol and drug dependence: matching techniques and populations. Washington, DC: U.S. Department of Health and Human Services; 1991.
132. Horvath A, Symonds B. Relation between working alliance and outcome in psychotherapy: a meta-analysis. *J Counsel Psychol.* 1991;38:139–149.
133. Connors GJ, Carroll KM, DiClemente CC et al. The therapeutic alliance and its relationship to alcoholism treatment participation and outcome. *J Consult Clin Psychol.* 1997;65:588–598.
134. Barber JP, Luborsky L, Gallop R, et al. Therapeutic alliance as a predictor of outcome and retention in the National Institute on Drug Abuse Collaborative Cocaine Treatment Study. *J Consult Clin Psychol.* 2001;69:119–124.
135. Luborsky L, Barber JP, Siqueland L. The revised Helping Alliance Questionnaire (HA-II): psychometric properties. *J Psychotherap Res.* 1996;5:260–271.
136. Petry NM, Bickel WK. Therapeutic alliance and psychiatric severity as predictors of completion of treatment for opioid dependence. *Psychiatr Serv.* 1999;50:219–227.
137. Shwartz M, Baker G, Mulvey KP, et al. Improving publicly funded substance abuse treatment: the value of case management. *Am J Public Health.* 1997;87:1659–1664.
138. Hilton ME, Maisto SA, Conigliaro J, et al. Improving alcoholism treatment across the spectrum of services. *Alcohol Clin Exp Res.* 2001;25:128–135.
139. Bachrach LL. Continuity of care for chronic mental patients: a conceptual analysis. *Am J Psychiatry.* 1981;138:1449–1456.
140. Baker F. Coordination of alcohol, drug abuse, and mental health services. Rockville, MD: U.S. Department of Health and Human Services; 1991.
141. Intagliata J. Improving the quality of community care for the chronically mentally disabled: the role of case management. *Schizophr Bull.* 1982;8:655–674.
142. Willenbring M. Psychiatric care management for chronic addictive disorders: conceptual framework. *Am J Addict.* 2001;10:242–248.
143. Laine C, Hauck WW, Gourevitch MN, et al. Regular outpatient medical and drug abuse care and subsequent hospitalization of persons who use illicit drugs. *JAMA.* 2001;285:2355–2362.
144. Des Jarlais DC, Fisher DG, Newman JC, et al. Providing hepatitis B vaccination to injection drug users: referral to health clinics vs on-site vaccination at a syringe exchange program. *Am J Public Health.* 2001;91:1791–1792.
145. Friedmann PD, Alexander JA, Jin L, et al. On-site primary care and mental health services in outpatient drug abuse treatment units. *J Behav Health Serv Res.* 1999;26:80–94.
146. Friedmann PD, D'Aunno TA, Jin L, et al. Medical and psychosocial services in drug abuse treatment: do stronger linkages promote client utilization? *Health Serv Res.* 2000;35:443–465.
147. Friedmann PD, Lemon SC, Stein MD, et al. Linkage to medical services in the Drug Abuse Treatment Outcome Study. *Med Care.* 2001;39:284–295.
148. Mezzelani P, Venturini L, Turrina G, et al. High compliance with a hepatitis B virus vaccination program among intravenous drug users. *J Infect Dis.* 1991;163:923.
149. Perlman DC, Perkins MP, Solomon N, et al. Tuberculosis screening at a syringe exchange program. *Am J Public Health.* 1997;87:862–863.
150. Stancliff S, Salomon N, Perlman DC, et al. Provision of influenza and pneumococcal vaccines to injection drug users at a syringe exchange. *J Subst Abuse Treat.* 2000;18:263–265.
151. Friedmann PD, Hendrickson JC, Gerstein DR, et al. Do mechanisms that link addiction treatment patients to primary care influence subse-

quent utilization of emergency and hospital care? *Med Care.* 2006;44: 8–15.
152. Friedmann P, Zhang Z, Hendrickson J, et al. Effects of primary medical care on addiction and medical severity in substance abuse treatment programs. *J Gen Intern Med.* 2003;18:1–8.
153. Magura S, Horgan CM, Mertens JR, et al. Effects of managed care on alcohol and other drug (AOD) treatment. *Alcohol Clin Exp Res.* 2002;26:416–422.
154. Saxon AJ, Malte CA, Sloan KL, et al. Randomized trial of onsite versus referral primary medical care for veterans in addictions treatment. *Med Care.* 2006;44:334–342.
155. O'Connor PG, Carroll KM, Shi JM, et al. Three methods of opioid detoxification in a primary care setting. A randomized trial. *Ann Intern Med.* 1997;127:526–530.
156. Basu S, Smith-Rohrberg D, Bruce RD, et al. Models for integrating buprenorphine therapy into the primary HIV care setting. *Clin Infect Dis.* 2006;42:716–721.
157. O'Connor PG, Farren CK, Rounsaville BJ, et al. A preliminary investigation of the management of alcohol dependence with naltrexone by primary care providers. *Am J Med.* 1997;103:477–482.
158. O'Connor PG, Oliveto AH, Shi JM, et al. A randomized trial of buprenorphine maintenance for heroin dependence in a primary care clinic for substance users versus a methadone clinic. *Am J Med.* 1998;105:100–105.
159. O'Connor PG, Waugh ME, Carroll KM, et al. Primary care-based ambulatory opioid detoxification: the results of a clinical trial. *J Gen Intern Med.* 1995;10:255–260.
160. O'Malley SS, Rounsaville BJ, Farren C, et al. Initial and maintenance naltrexone treatment for alcohol dependence using primary care vs specialty care: a nested sequence of 3 randomized trials. *Arch Intern Med.* 2003;163:1695–1704.
161. Kiritze-Topor P, Huas D, Rosenzweig C, et al. A pragmatic trial of acamprosate in the treatment of alcohol dependence in primary care. *Alcohol Alcohol.* 2004;39:520–527.
162. Weisner C, Mertens J, Parthasarathy S, et al. Integrating primary medical care with addiction treatment: a randomized controlled trial. *JAMA.* 2001;286:1715–1723.
163. Willenbring M, Olson D, Bielinski J. Integrated outpatients treatment for medically ill alcoholic men: results from a quasi-experimental study. *J Stud Alcohol.* 1995;56:337–343.
164. Willenbring M, Olson D. A randomized trial of integrated outpatient treatment for medically ill alcoholic men. *Arch Intern Med.* 1999;159: 1946–1952.
165. Lieber CS, Weiss DG, Groszmann R, et al. I. Veterans Affairs Cooperative Study of polyenylphosphatidylcholine in alcoholic liver disease: effects on drinking behavior by nurse/physician teams. *Alcohol Clin Exp Res.* 2003;27:1757–1764.
166. Andersen M, Paliwoda J, Kaczynski R, et al. Integrating medical and substance abuse treatment for addicts living with HIV/AIDS: evidence-based nursing practice model. *Am J Drug Alcohol Abuse.* 2003;29:847–859.
167. Bartels SJ, Coakley EH, Zubritsky C, et al. Improving access to geriatric mental health services: a randomized trial comparing treatment engagement with integrated versus enhanced referral care for depression, anxiety, and at-risk alcohol use. *Am J Psychiatry.* 2004; 161:1455–1462.
168. World Health Organization. Executive summary and global report. Innovative care for chronic conditions: building blocks for action. Noncommunicable diseases and mental health. Available at: http://www.who.int/chp/en/. Accessed July 17, 2007.
169. Casalino L, Gillies RR, Shortell SM, et al. External incentives, information technology, and organized processes to improve health care quality for patients with chronic diseases. *JAMA.* 2003;289:434–441.
170. Institute of Medicine. Improving the Quality of Health Care for Mental and Substance-Use Conditions: Quality Chasm Series. Washington, DC: The National Academies Press; 2005.
171. Samet JH, Larson M, Horton NJ, et al. Linking alcohol and drug dependent adults to primary medical care: a randomized controlled trial of a multidisciplinary health intervention in a detoxification unit. *Addiction.* 2003;98:509–516.

Alcohol Counseling Reflects Higher Quality of Primary Care

Richard Saitz, MD, MPH[1,2,3], Nicholas J. Horton, ScD[6], Debbie M. Cheng, ScD[1,4], and Jeffrey H. Samet, MD, MA, MPH[1,5]

[1]Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Boston Medical Center and Boston University School of Medicine, Boston, MA, USA; [2]Youth Alcohol Prevention Center, Boston University School of Public Health, Boston, MA, USA; [3]Department of Epidemiology, Boston University School of Public Health, Boston, MA, USA; [4]Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA; [5]Department of Social and Behavioral Sciences, Boston University School of Public Health, Boston, MA, USA; [6]Department of Mathematics and Statistics, Smith College, Northampton, MA, USA.

BACKGROUND: Some primary care physicians do not conduct alcohol screening because they assume their patients do not want to discuss alcohol use.

OBJECTIVES: To assess whether (1) alcohol counseling can improve patient-perceived quality of primary care, and (2) higher quality of primary care is associated with subsequent decreased alcohol consumption.

DESIGN: A prospective cohort study.

SUBJECTS: Two hundred eighty-eight patients in an academic primary care practice who had unhealthy alcohol use.

MEASUREMENTS: The primary outcome was quality of care received [measured with the communication, whole-person knowledge, and trust scales of the Primary Care Assessment Survey (PCAS)]. The secondary outcome was drinking risky amounts in the past 30 days (measured with the Timeline Followback method).

RESULTS: Alcohol counseling was significantly associated with higher quality of primary care in the areas of communication (adjusted mean PCAS scale scores: 85 vs. 76) and whole-person knowledge (67 vs. 59). The quality of primary care was not associated with drinking risky amounts 6 months later.

CONCLUSIONS: Although quality of primary care may not necessarily affect drinking, brief counseling for unhealthy alcohol use may enhance the quality of primary care.

KEY WORDS: alcohol; counseling; brief intervention; quality of primary care.
J Gen Intern Med 23(9):1482–6
DOI: 10.1007/s11606-008-0574-4

Results of this study were presented at the following meetings: the annual national meeting of the Society of General Internal Medicine, Chicago, May 2004 and the annual national meeting of the Research Society on Alcoholism, Vancouver, Canada, June 2004. This study was funded by a grant from the Robert Wood Johnson Foundation (grant 031489). Dr. Samet received support from the National Institute on Alcohol Abuse and Alcoholism (K24-AA015674).
Received May 22, 2007
Revised October 29, 2007
Accepted February 25, 2008
Published online July 10, 2008

BACKGROUND

Practice guidelines recommend that clinicians screen and offer brief intervention for unhealthy alcohol use (the spectrum from drinking at-risk amounts through dependence) in adults[1,2]. Despite these guidelines and available efficacious strategies, unhealthy alcohol use among primary care patients is often unrecognized[3,4] and treated ineffectively[5,6].

Many barriers to addressing unhealthy alcohol use exist, including the assumption held by some physicians that patients do not want to discuss drinking. Physicians who are concerned about alienating their patients or believe their patients lack interest in discussing alcohol use will either avoid raising the subject or may not address it adequately[7,8]. These doctors may also worry that alcohol counseling will diminish patient-perceived quality of care[9].

Most patients, however, are not bothered by alcohol discussions and may welcome them[10,11]. They often find the discussions useful[3] and are more likely to be satisfied with their care than are patients who do not have such discussions[12].

Still, whether alcohol counseling is associated with higher quality of care remains unknown. Therefore, we conducted this study of patients with unhealthy alcohol use to determine whether alcohol counseling during a primary care visit influences patient-perceived quality of primary care. Further, we studied whether quality of care is associated with drinking of risky amounts.

METHODS

Subjects

Subjects were patients in an urban, academic primary care practice who had participated in a randomized trial testing the effects of providing physicians with patients' alcohol screening results[5]. In that cluster randomized trial, physicians were randomly assigned to receive or not receive the results of alcohol screening that was done in the waiting room prior to the physician visit. Patients had unhealthy alcohol use and presented for a visit with the physician and were identified in the waiting room by screening. The intervention consisted of a sheet of paper summarizing the results of the CAGE test, recent drinking amounts, and readiness to change. Eligible subjects spoke English or Spanish, drank in the past month, and had either a ≥ 1 on the CAGE alcohol screening test[13] or drank risky amounts (past 30 days; Table 1)[14].

Table 1. Characteristics at Enrollment: 288 Subjects with Unhealthy Alcohol Use

	Counseled[a] about drinking n=132	Not counseled about drinking n=156	P value
Male, no. (%)	87 (66)	90 (58)	0.15
Age, mean (SD)	45 (13)	41 (13)	0.009
Race/ethnicity			0.74
African American, no. (%)	80 (61)	85 (54)	
White, no. (%)	22 (17)	32 (21)	
Latino, no. (%)	20 (15)	25 (16)	
High school education, no. (%)	64 (48)	115 (74)	<0.001
Medical comorbidity,[b] ever, no. (%)	91 (69)	107 (69)	0.95
Drinks per drinking day,[c] past 30 days, mean (SD)	7 (5)	5 (4)	0.002
Alcohol problems,[d] current, mean score (SD)	11 (11)	5 (9)	<0.001
Drank risky amounts,[e] past 30 days, no. (%)	108 (82)	113 (72)	0.06
Readiness to change,[f] mean (SD)	5.8 (3.0)	4.9 (3.3)	0.02
Met physician previously, no. (%)	96 (73)	109 (70)	0.59
Wanted the physician they were seeing to provide general information about alcohol use, no. (%)	78 (59)	76 (49)	0.08
Wanted the physician they were seeing to give advice about their drinking habits,[g] no. (%)	83 (63)	82 (53)	0.09
Had a physician who was randomized to the intervention group in the randomized controlled trial, no. (%)	72 (55)	80 (51)	0.58
Had a physician who was faculty, no. (%)	106 (80)	116 (74)	0.23

[a]*Based on patient self-report*
[b]*Determined with the method of Katz et al* [16]
[c]*Determined by the Timeline Followback method, which assesses the type and number of standard drinks consumed on each of the previous 30 days* [19]
[d]*Short Inventory of Problems (SIP 2R) total score* [20]
[e]*>14 standard drinks per week or >4 drinks per occasion for men; >7 drinks per week or >3 drinks per occasion for women and people ≥66 years* [14]
[f]*Based on a visual analogue scale ranging from 0 to 10* [15]*; n=114 for counseled, 149 for not counseled*
[g]*n=155 for not counseled*

Enrolled subjects provided written informed consent and were compensated. The Institutional Review Board at Boston Medical Center approved this study.

Measurements

Research associates (RAs) screened patients waiting to see one of 40 primary care physicians, for eligibility through a self-administered questionnaire (there was no other basis for selection). RAs then interviewed enrolled subjects immediately before and immediately after the physician visits.

During the interview before the visit, RAs assessed readiness to change (visual analogue scale from 0 to 10)[15] and medical comorbidity[16]. Immediately after the visit, RAs asked patients whether they had received alcohol counseling (a referral and/or advice on safe drinking limits, decreasing intake, or abstaining) during the visit and about quality of care based on three (of 11) scales from the Primary Care Assessment Survey (PCAS)[17], a validated tool that measures the fundamental characteristics of primary care defined by the Institute of Medicine[18]. The scales, ranging from 0 to 100 with 100 indicating the highest level of performance, included communication (e.g., attention to what patients say); whole-person knowledge (e.g., physician's knowledge of a patient's health concerns, values, and beliefs); and trust (e.g., physicians' integrity). Lastly, RAs evaluated subjects' alcohol consumption (past 30 days, Timeline Followback method)[19] and current alcohol problems [Short Inventory of Problems (SIP 2R)][20]. Six months later, RAs interviewed subjects by telephone.

Outcomes

The primary outcome was patient-perceived primary care quality, measured with the three PCAS scales[17] immediately after physician visits. The secondary outcome was drinking risky amounts at the 6-month follow-up.

Statistical Analyses

We performed all analyses using SAS software, version 8.1. (SAS Institute, Cary, North Carolina). We used the chi-square test and t test, as appropriate, for bivariate comparisons. Reported P values are two-tailed; a P value of <0.05 was considered statistically significant.

We used linear mixed effects models to test the association between alcohol counseling and the three PCAS scales and generalized estimating equations (GEE) logistic regression models to test the associations between the PCAS scales and drinking. These correlated data models were used to adjust for clustering of patients by physician (exchangeable working correlations 0.03 to 0.08 for PCAS scales). The mixed model used an exchangeable working covariance structure and the GEE model used an independence working correlation structure.

RESULTS

Of eligible patients, 55% enrolled (Fig. 1). Enrolled and eligible but not enrolled subjects were similar on age, sex, race, and CAGE questionnaire responses; however, enrolled subjects

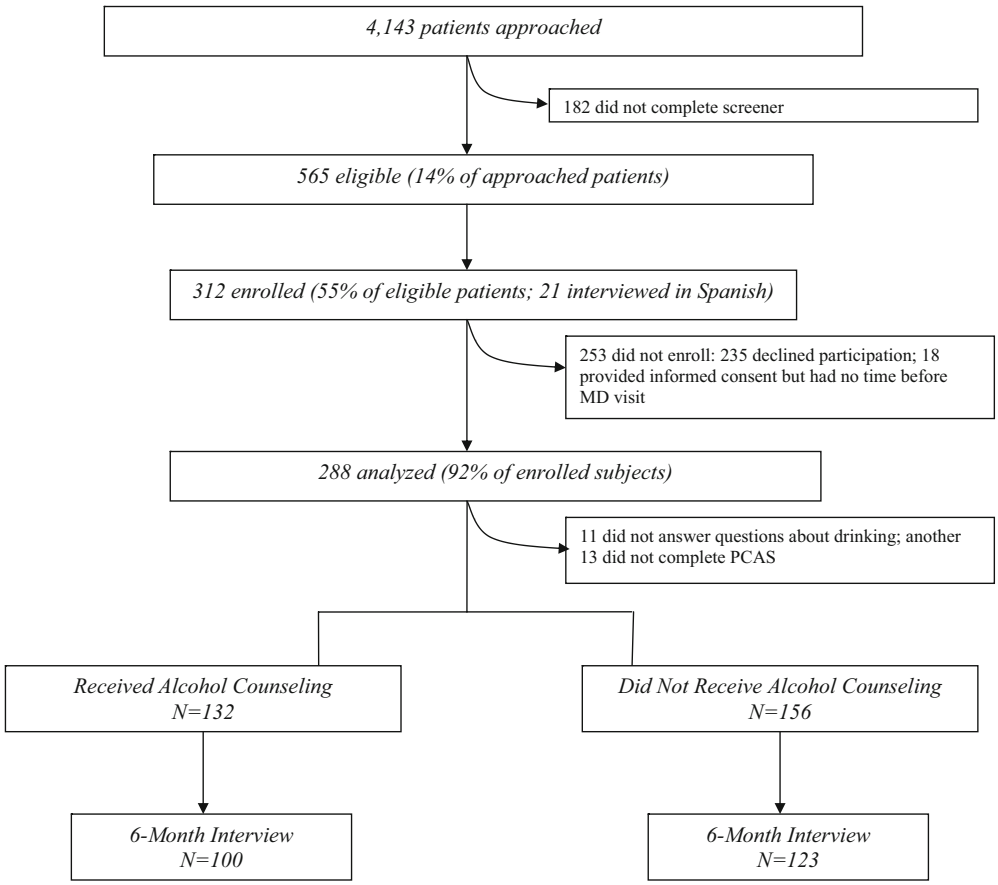


Fig. 1. Subject Enrollment and Follow-up.

had significantly greater alcohol consumption (drinks/drinking day, 4.5 vs. 3.4) and readiness to change their drinking (mean score, 5.5 vs. 4.9).

Of the enrolled sample, 301 (96%) answered questions about alcohol counseling during their primary care visit; 288 (96%) of these completed the PCAS and compose our sample. At 6 months, 223 of the 288 (77%) were assessed. Compared to those lost to follow-up, interviewed subjects were significantly more likely to have a high school education (68% vs. 41%) and to have met their doctor before (74% vs. 60%).

Mean (SD) PCAS scores were communication 81 (SD 16), comprehensiveness 66 (SD 21), and trust 80 (SD 12). Almost half of the sample [132 (46%)] reported receiving alcohol counseling during their primary care visit. Counseled subjects were significantly more likely than subjects who had not been counseled to be older, have no high school education, and have a higher mean number of drinks/drinking day, alcohol problem score, and readiness to change (Table 1).

In unadjusted analyses, counseled subjects reported higher quality of primary care in the areas of communication, whole-person knowledge, and trust, though the latter was not statistically significant (Table 2). These findings persisted in multivariable analyses.

At 6 months, 121 of 223 subjects (54%) were drinking risky amounts. Quality of primary care did not significantly affect the odds of drinking risky amounts [adjusted odds ratios, 1.0

Table 2. Alcohol Counseling and Quality of Primary Care

Primary care quality domain	Unadjusted mean scores[a] (95% CI)			Adjusted mean scores[b] (95% CI)		
	Counseled *n*=132	Not counseled *n*=156	*P* value	Counseled *n*=132	Not counseled *n*=156	*P* value
Communication	85 (83–87)	78 (75–80)	<0.001	85 (81–88)	76 (73–79)	<0.001
Whole-person knowledge	70 (67–74)	62 (59–65)	0.001	67 (62–71)	59 (55–63)	0.005
Trust	81 (79–83)	79 (77–80)	0.08	79 (77–82)	77 (74–79)	0.06

[a]*Unadjusted analyses account for clustering of patients by physician.*
[b]*Adjusted for sex, race, education, comorbidity, randomization, level of physician training, having met the physician previously, mean drinks per drinking day, alcohol problem score, and clustering of patients by physician*

(95% CI, 0.98–1.02) for communication; 1.00 (95% CI, 0.98–1.01) for whole-person knowledge; and 1.00 (95% CI, 0.98–1.03) for trust].

CONCLUSION

Alcohol counseling by primary care physicians was associated with higher patient-perceived quality of care, specifically better communication, and whole-person knowledge. Higher quality of care, however, was not associated with decreased drinking of risky amounts at 6 months.

This study is novel as it assesses the relationships between (1) alcohol counseling and quality of primary care with a validated measure and (2) quality of primary care and drinking outcomes. Our study supports results from previous research indicating that patients are not bothered by, and often appreciate, being asked during primary care visits about their alcohol use[3,10–12]. The magnitude of differences in quality we observed was similar to, though generally smaller than, those known to impact clinical outcomes[17,21]. For example, Kim et al reported that single standard deviation increases in primary care quality were associated with a lower risk of subsequent substance use[21]. While various studies have reported a link between primary care quality and health outcomes[9,21], ours did not. High-quality primary care may be necessary, but not sufficient, to help patients reduce their drinking. The lack of association between quality of primary care and decreased consumption is most likely because specific elements of brief interventions that are essential to change drinking (e.g., targeted advice) were not offered in this study.

Our study has several strengths. We used a standard measure of drinking in a sample with a range of unhealthy alcohol use and a well-validated measure of primary care quality that has been linked to clinical outcomes. The PCAS and its individual subscales have high internal consistency and reliability; each subscale has been validated[17]. Lastly, we used a prospective design and assessed counseling and quality immediately after a primary care visit.

Several limitations should be considered. First, we could not determine whether alcohol counseling affects quality beyond the self-report measures assessed. However, the measures we chose are among the best ways to assess primary care quality and are particularly relevant to alcohol counseling[17]. Second, we assessed the drinking outcome at only one timepoint. This method is similar to that used in studies supporting brief intervention for unhealthy alcohol use[2]. Third, because this was an observational study, our ability to determine causality is limited; however, we did adjust analyses for potential confounding factors. Fourth, the initial research assessment may have sensitized subjects and influenced their responses to questions about perceived quality. Fifth, most subjects had visited their physicians and discussed alcohol previously. Therefore, the observed associations between counseling and quality of care may be biased towards the null; nonetheless, we observed some effects. Sixth, intervening influences (e.g., participation in Alcoholics Anonymous) could have affected drinking outcomes. Brief counseling, however, is known to reduce consumption beyond such influences. Lastly, the differences between the enrolled and nonenrolled patients limited generalizability and, along with the differences in those followed and lost to follow-up, may have biased analyses (the latter limited to the drinking analyses). However, the direction of bias resulting from these differences is difficult to predict.

Physicians should conduct alcohol counseling for unhealthy alcohol use for many reasons. Alcohol counseling has proven efficacy in outpatient settings and is recommended in practice guidelines. Furthermore, most patients want to receive advice about their drinking, and as indicated by this study, such a discussion does not diminish quality of care. These findings provide evidence that screening and intervention for unhealthy alcohol use may improve quality of care from the patient's perspective.

ACKNOWLEDGEMENT: *The authors are indebted to Rosanne T. Guerriero, MPH, for her efforts in manuscript writing and editing. This study was funded by a grant from the Robert Wood Johnson Foundation (grant 031489). Dr. Samet received support from the National Institute on Alcohol Abuse and Alcoholism (K24-AA015674).*

Conflict of Interest: *None disclosed.*

Corresponding Author: *Richard Saitz, MD, MPH; Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Boston Medical Center and Boston University School of Medicine, 801 Massachusetts Avenue, 2nd floor, Boston, MA 02118, USA (e-mail: rsaitz@bu.edu).*

REFERENCES

1. **Saitz R**. Clinical practice. Unhealthy alcohol use. N Engl J Med. 2005;352(6):596–607.
2. US Preventive Services Task Force. Screening and behavioral counseling interventions in primary care to reduce alcohol misuse: recommendation statement. Ann Intern Med. 2004;140(7):554–6.
3. **Aalto M, Seppa K**. Usefulness, length and content of alcohol-related discussions in primary health care: the exit poll survey. Alcohol Alcohol. 2004;39(6):532–5.
4. **Bradley KA, Curry SJ, Koepsell TD, Larson EB**. Primary and secondary prevention of alcohol problems: US internist attitudes and practices. J Gen Intern Med. 1995;10(2):67–72.
5. **Saitz R, Horton NJ, Sullivan LM, Moskowitz MA, Samet JH**. Addressing alcohol problems in primary care: a cluster randomized, controlled trial of a systems intervention. The screening and intervention in primary care (SIP) study. Ann Intern Med. 2003;138(5):372–82.
6. **McGlynn E, Asch S, Adams J, et al**. The quality of health care delivered to adults in the United States. N Engl J Med. 2003;348(26):2635–45.
7. **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(2):84–91.
8. **Beich A, Gannik D, Malterud K**. Screening and brief intervention for excessive alcohol use: qualitative interview study of the experiences of general practitioners. BMJ. 2002;325(7369):870–4.
9. **Safran DG, Taira DA, Rogers WH, Kosinski M, Ware JE, Tarlov AR**. Linking primary care performance to outcomes of care. J Fam Pract. 1998;47(3):213–20.
10. **Aalto M, Pekuri P, Seppa K**. Primary health care professionals' activity in intervening in patients' alcohol drinking: a patient perspective. Drug Alcohol Depend. 2002;66(1):39–43.
11. **Zimmerman M, Farber NJ, Hartung J, Lush DT, Kuzma MA**. Screening for psychiatric disorders in medical patients: a feasibility and patient acceptance study. Med Care. 1994;32(6):603–8.
12. **Steven ID, Thomas SA, Eckerman E, Browning C, Dickens E**. The provision of preventive care by general practitioners measured by patient completed questionnaires. J Qual Clin Pract. 1999;19(4):195–201.
13. **Mayfield D, McLeod G, Hall P**. The CAGE questionnaire: validation of a new alcoholism screening instrument. Am J Psychiatry. 1974;131(10):1121–3.

14. **Saunders JB, Lee NK**. Hazardous alcohol use: its delineation as a subthreshold disorder, and approaches to its diagnosis and management. Compr Psychiatry. 2000;41(2 Suppl 1):95–103.
15. **Rollnick S, Heather N, Gold , Hall W**. Development of a short 'readiness to change' questionnaire for use in brief, opportunistic interventions among excessive drinkers. Br J Addict. 1992;87(5):743–54.
16. **Katz JN, Chang LC, Sangha O, Fossel AH, Bates DW**. Can comorbidity be measured by questionnaire rather than medical record review? Med Care. 1996;34(1):73–84.
17. **Safran DG, Kosinski M, Tarlov A, et al**. The Primary Care Assessment Survey: tests of data quality and measurement performance. Med Care. 1998;36(5):728–39.
18. Institute of Medicine. Primary care: America's health in a new era. Washington, D.C.: National Academy Press; 1996.
19. **Sobell LC, Brown J, Leo GI, Sobell MB**. The reliability of the Alcohol Timeline Followback when administered by telephone and by computer. Drug Alcohol Depend. 1996;42(1):49–54.
20. **Miller W, Tonigan J, Longabaugh R**. The Drinker Inventory of Consequences (DrInC). An Instrument for Assessing Adverse Consequences of Alcohol Abuse. Test Manual. Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism; 1995.
21. **Kim TW, Samet JH, Cheng DM, Winter M, Safran DG, Saitz R**. Primary care quality and addiction severity: a prospective cohort study. Health Serv Res. 2007;42(2):755–72.

Addiction

RESEARCH REPORT

doi:10.1111/j.1360-0443.2008.02251.x

Mitigating risky sexual behaviors among Russian narcology hospital patients: the PREVENT (Partnership to Reduce the Epidemic Via Engagement in Narcology Treatment) randomized controlled trial

Jeffrey H. Samet[1,2], Evgeny M. Krupitsky[3], Debbie M. Cheng[4], Anita Raj[2], Valentina Y. Egorova[3], Suzette Levenson[5], Seville Meli[1], Carly Bridden[1], Elena V. Verbitskaya[3], Mary L. Kamb[6] & Edwin E. Zvartau[3]

Boston University School of Medicine/Boston Medical Center, Department of Medicine, Section of General Internal Medicine, Clinical Addiction Research and Education (CARE) Unit, USA,[1] Boston University School of Public Health, Department of Social and Behavioral Sciences, MA, USA,[2] St Petersburg State Pavlov Medical University, St Petersburg, Russian Federation,[3] Boston University School of Public Health, Department of Biostatistics, MA, USA,[4] Boston University School of Public Health, Data Coordinating Center, MA, USA[5] and Centers for Disease Control and Prevention (CDC), Division of STD Prevention, Atlanta, GA, USA[6]

ABSTRACT

Aim To assess the effectiveness of a sexual risk reduction intervention in the Russian narcology hospital setting. **Design, setting and participants** This was a randomized controlled trial from October 2004 to December 2005 among patients with alcohol and/or heroin dependence from two narcology hospitals in St Petersburg, Russia. **Intervention** Intervention subjects received two personalized sexual behavior counseling sessions plus three telephone booster sessions. Control subjects received usual addiction treatment, which did not include sexual behavior counseling. All received a research assessment and condoms at baseline. **Measurements** Primary outcomes were percentage of safe sex episodes (number of times condoms were used ÷ by number of sexual episodes) and no unprotected sex (100% condom use or abstinence) during the previous 3 months, assessed at 6 months. **Findings** Intervention subjects reported higher median percentage of safe sex episodes (unadjusted median difference 12.7%; $P = 0.01$; adjusted median difference 23%, $P = 0.07$); a significant difference was not detected for the outcome no unprotected sex in the past 3 months [unadjusted odds ratio (OR) 1.6, 95% confidence interval (CI) 0.8–3.1; adjusted OR 1.5, 95% CI 0.7–3.3]. **Conclusions** Among Russian substance-dependent individuals, sexual behavior counseling during addiction treatment should be considered as one potential component of efforts to decrease risky sexual behaviors in this HIV at-risk population.

Keywords Alcohol, behavioral intervention, heroin, HIV prevention, narcology hospital, randomized controlled trial, Russia, sex risk behaviors, substance dependence.

Correspondence to: Jeffrey H. Samet, Section of General Internal Medicine, Boston Medical Center, 801 Massachusetts Avenue, 2nd Floor, Boston, MA 02118, USA. E-mail: jsamet@bu.edu

Submitted 10 July 2007; initial review completed 6 November 2007; final version accepted 12 March 2008

INTRODUCTION

Russia has one of the fastest-growing acquired immune deficiency syndrome (AIDS) epidemics in the world, with an estimated 860 000 human immunodeficiency virus (HIV)-infected people in 2003 [1]. In St Petersburg, the prevalence of HIV increased 100-fold (0.013–1.3%) from 1998 to 2002 [2,3]. Initially the Russian HIV epidemic was almost exclusively among injection drug users (IDUs) [4]; however, concern exists that HIV is expanding into the general population via sexual transmission [1,2,5]. Among sex workers, HIV seroprevalence is estimated at 5–15% overall, but 48% in those who also inject drugs [2].

Alcohol use, highly prevalent in Russia [6,7], may increase high-risk sexual behaviors (e.g. multiple sex partners, unprotected sex) among IDUs and alcohol-dependent individuals [8–11]. Among female drug users, increased alcohol consumption has been associated with sexual HIV risk-taking behavior [12]. Furthermore,

animal models suggest that alcohol consumption plays a permissive role for HIV replication, as the resultant higher viral loads may increase risk of transmission [13,14].

In addition to the association between alcohol use and high-risk sexual behavior, extensive evidence demonstrates that users of other substances may also be commonly involved in high-risk sex [15–19]. Use of stimulants is associated with increased sexual risk behavior such as unprotected sex, multiple partners and selling sex in the United States and Russia [16,17,19]. Studies with Russian IDUs have found that multiple partners, having an IDU sex partner and unprotected sex, particularly with steady sex partners, is common [18,20–22]. Selling sex for drugs, money, goods or shelter is also reported commonly among IDUs in St Petersburg [16,18].

Behavioral interventions to reduce risky sex are an essential component to HIV prevention and are even more critical in the absence of a cure or vaccine. Interventions that focus on personal risk reduction have been shown to be effective in reducing sexual risk behaviors and diseases in developed as well as developing countries [23–25]. A study of meta-analyses of HIV prevention interventions showed that programs targeting drug users can be successful at reducing sexual risk [26]. Several moderators were identified as contributing to intervention efficacy, including separate gender sessions, didactic lecture, self-control/coping skills and greater number of intervention techniques used. Although sex risk behaviors can be different based on HIV serostatus [27,28], few data have been reported that describe whether response to a particular sexual risk reduction intervention depends upon HIV status. Successful prevention interventions among HIV-positive individuals share many of the following characteristics: based on behavioral theory, targeted HIV transmission behaviors, delivered by health-care providers or counselors to individuals, time-intensive, delivered in a familiar medical or service environment, provided skills building or addressed other HIV-related issues [24]. As yet, relatively few controlled trials of behavioral interventions have demonstrated efficacy in reducing sexual HIV risk among substance users in treatment settings [29–31]. Although many of the individual studies included in a meta-analysis on this topic had positive effect estimates, most failed to reach statistical significance on their own [32].

Current HIV prevention efforts in Russia address sexual risk reduction mainly in mass media promotion of condom use and encouragement of HIV counseling and testing [33–35]. A limited number of non-governmental organizations disseminate information on HIV prevention among IDUs [35]. Treatment of opioid dependence with methadone or buprenorphine, proven effective at preventing HIV among IDUs [36], is illegal in Russia [35,37]. Regional narcology hospitals play a central role in Russia's efforts to address alcohol and drug dependence but have not addressed HIV aggressively.

Reducing risky sexual behaviors among alcohol- and drug-dependent individuals is an HIV intervention strategy that has not, as yet, been pursued in Russia [16,38]. We tested such an intervention among narcology patients in St Petersburg, Russia.

METHODS

Study design

The Russian PREVENT (Partnership to Reduce the Epidemic Via Engagement in Narcology Treatment) study was a randomized controlled trial (RCT) [39] that recruited men and women with alcohol and/or drug dependence from two substance abuse treatment facilities near St Petersburg, Russia [i.e. Leningrad Regional Center for Addictions (LRCA) and the Medical Narcology Rehabilitation Center (MNRC)]. Narcology hospitals are a standard treatment setting for drug- and alcohol-dependent individuals in eastern Europe. Hospitalization is typically 3–4 weeks, in which initial addiction treatment follows detoxification.

Trained physician research associates approached patients after initial detoxification, assessed eligibility, offered participation and conducted assessments. Eligibility criteria were the following: age 18 years and older; a primary diagnosis of alcohol or drug dependence; no alcohol or other abused substances for at least 48 hours; reported unprotected anal or vaginal sex in the past 6 months; willingness to undergo HIV testing per standard narcology hospital protocols or previous diagnosis of HIV infection; and provision of reliable contact information (i.e. a home telephone number, an address within 150 km of St Petersburg and one friend or family contact). Patients not fluent in Russian or with cognitive impairment based on the research associates' judgement were excluded. Participants provided written informed consent prior to enrollment in the study. The Institutional Review Boards of Boston Medical Center and St Petersburg Pavlov State Medical University approved this study.

Subject assessment

Baseline assessments occurred after randomization; however, subjects and assessors were blinded to intervention group at this point. Follow-up assessments occurred 3 months and 6 months after enrollment. All follow-up assessments were conducted with patients after discharge from the hospital. Assessment data included demographics, behavioral intentions for condom and needle use, Center for Epidemiologic studies Depression (CES-D) Scale

for depressive symptoms [40], history of sexually transmitted diseases, HIV testing and disclosure, ICD-10 substance dependence diagnosis and the Short Form 36 (SF-36) General Health Survey [41]. Questions about HIV sex and drug risk behaviors came from multiple sources: the RESPECT study [42] (e.g. 'in the past 3 months, how many times have you had vaginal sex with your primary partner; how many of those times did you use a condom?'); the Risk Assessment Battery (RAB) [43] sex and drug use subscales (e.g. 'in the past 6 months, how often were you paid to have sex?'); the Timeline Follow-back survey (TLFB) [44,45] (e.g. total number of drinks on each day in the past 30 days, number of times condoms used with vaginal sex on each day in the past 30 days); and the Addiction Severity Index-*Lite* [46]. All instruments were translated from English to Russian for this study (e.g. RESPECT and TLFB), unless already available in Russian (e.g. RAB). Risk behaviors were assessed by both face-to-face interviews and to promote truth-telling through an Audio Computer-Assisted Self Interviewing (ACASI) system at baseline and 6 months. The 3-month assessment, administered via the telephone, included RESPECT and RAB questions about HIV risk behaviors. The 6-month assessment was identical to the baseline and occurred at the narcology hospitals within a 5–7-month window after enrollment. All interviews were conducted in Russian by trained personnel not involved in interventions and who were blinded to treatment group. Subjects were compensated the equivalent of US$5, $5 and $30 for the baseline and 3- and 6-month assessments, respectively, and all received 30 condoms at baseline.

Study treatments

Subjects were assigned randomly to either the Russian PREVENT program (intervention group) or standard addiction treatment (control group).

Russian PREVENT program

The Russian PREVENT intervention was based on the Brief Counseling model used in Project RESPECT, a prevention program tested in US sexually transmitted disease (STD) clinics, which demonstrated reduction in risky sexual behaviors and STDs [42]. RESPECT involved a two-session HIV prevention counseling intervention used with HIV testing to increase participants' perception of personal risk, support participant-initiated changes and identify small, achievable steps towards reducing personal risk. The Russian PREVENT study was modified from Project RESPECT by the US–Russian team to meet the needs of the Russian narcology hospital setting and patients. Modifications included the following: (1) enrollment of known HIV-positive as well as negative participants (all patients were HIV tested as part of the narcology program); (2) emphasis on basic HIV prevention and transmission knowledge due to relatively low access to such information among this St Petersburg population; (3) inclusion of booster telephone sessions to sustain programmatic effects by providing longer-term support; and (4) provision of skills-building on HIV-related risk reduction for sexual and injection drug use behaviors for both HIV-infected and uninfected individuals. The modifications yielded longer PREVENT sessions than RESPECT (30–60 versus 20 minutes). Sessions occurred at the narcology hospitals and involved provision of HIV test results, discussion of personal risk and risk reduction and creation of a behavioral change plan. The first session included a personal assessment of HIV risk, discussion of HIV risk perceptions and negotiation of a personalized risk reduction plan. The second session was held within 1 week of session 1 to allow sufficient time for HIV test results to return. The interventionist provided the HIV test results and reviewed the risk reduction plan (i.e. promotion of safer sex via condom skills, sexual-negotiation skills building, developing positive attitudes regarding safer sex and emphasizing the role of alcohol). Additional content was covered as appropriate for HIV-infected subjects (e.g. HIV disclosure) and injection drug users (e.g. clean needle use). The same interventionist delivered both intervention sessions to an individual. Booster sessions after hospital discharge occurred via telephone monthly for 3 months, when interventionists checked in and updated participants' personal long-term risk reduction goals and plans. Typically, the same interventionist delivered the in-hospital and booster call sessions, but for a minority (approximately 5%) another interventionist conducted the booster calls.

Control group program

Subjects randomized to the control condition received usual addiction treatment at the narcology hospital, including HIV testing, but no sexual behavior counseling. Those known to be HIV-infected or who tested positive received one 20-minute HIV post-test counseling session with the study interventionists, even though this was not standard care in the narcology hospitals. This counseling for these control individuals with HIV infection included creation of risk reduction goals and referral to an HIV care program. All control subjects were contacted for study checks, but not counseled, at the booster time-points. Both control and intervention subjects received 30 condoms.

Training of interventionists

Interventionists (two psychiatrists and a psychologist trained in HIV and addiction) were trained by US

collaborators with HIV and substance use intervention experience (A. R., J. H. S.); they were trained about both general risk reduction interview techniques and the Russia-adapted RESPECT intervention. The lead interventionist (V. E.) underwent an initial training in English in the United States. A subsequent 3-day training in St Petersburg with all interventionists using simultaneous translation allowed multiple role-playing sessions to be observed and critiqued by the behavioral psychologist (A. R.).

Quality assurance procedures

The following efforts were conducted to ensure fidelity of the Russian PREVENT intervention: (1) 20% of each interventionist's subjects were selected randomly to have their sessions observed by another interventionist. The observer documented whether the curriculum content and activities were covered and how well the interventionists achieved session objectives. All observed sessions demonstrated 100% coverage of the curriculum material and activities. In 90% of observed cases, interventionists were described as implementing the program at an 'excellent' level in a variety of domains (e.g. providing HIV risk assessment and counseling, establishing rapport). (2) Interventionists participated in monthly research team meetings to discuss programmatic difficulties; no major problems with fidelity were noted, although there were difficulties in completing the booster session observations. (3) Participants completed a brief survey to assess perceptions of the utility of the program in helping to reduce their HIV/STD risk, the competence of the interventionist and whether they would recommend the program to others at the hospital. All responded that they found the program somewhat or very informative and helpful for reducing their HIV risk and in answering questions about HIV and that they would recommend the program to other patients at the hospital. The research team members discussed the results of these quality assurance assessments regularly.

Primary outcomes

The two primary outcomes of interest were (i) percentage of safe sex episodes and (ii) no unprotected sex (yes/no) during the past 3 months. These outcomes were assessed at the 6-month follow-up visit by ACASI. Percentage of safe sex episodes, a continuous variable, was defined as the percentage of times condoms were used out of the total number of sexual episodes (anal and vaginal intercourse) in the past 3 months. No unprotected sex was defined as either 100% condom use during anal and vaginal intercourse or sexual abstinence.

Secondary outcomes

Secondary outcomes included (i) number of unsafe sex episodes (i.e. no condom use during anal or vaginal sex; people who did not have sex were coded as having no unsafe sex episodes); and (ii) any condom use during the past 3 months by ACASI at the 6-month follow-up. Sex risk behaviors at the 3-month telephone follow-up were also examined: percentage of safe sex episodes, no unprotected sex and any condom use.

Randomization and blinding

Random allocation of subjects was accomplished using a computer-generated list of random numbers using permuted blocks stratified according to gender and dependence diagnosis. Research associates assigned subjects to the intervention or control condition immediately after completion of informed consent. Three strata of dependence diagnosis were used: alcohol, drug or dual (alcohol and drug). The research associate who assessed outcomes and contacted subjects to arrange follow-up appointments remained blinded to treatment assignments throughout follow-up.

Statistical analysis

Analyses were conducted using logistic regression and median regression models to adjust for possible group differences at baseline. All outcome analyses used the intention-to-treat principle, two-sided tests and a significance level of 0.05, with SAS/STAT software version 9.1. The observed sample sizes at 6 months allowed us 80% power to detect a minimum 25% difference (assuming the observed proportion of 29% for control group) in no unprotected sex using a two-sided χ^2 test with continuity correction.

Subject enrollment and baseline characteristics

Of the 329 patients screened, 181 met inclusion criteria and provided consent to participate. Eighty-seven subjects were randomized to the control and 94 to the intervention treatment group. Participants were 75% male, median age 30 years and 15% HIV-infected. Dependence diagnoses were 60% alcohol, 32% heroin and 8% dual. Among alcohol-dependent and dual-diagnosis subjects ($n = 123$), the median number of drinks per day reported at baseline was 5.4 [interquartile range (IQR): 2.6–10.0]. Among heroin-dependent and dual-diagnosis subjects ($n = 73$), 97% reported injection drug use in the past 6 months. At baseline, the two groups were similar in all examined characteristics except 'reported buying or selling sex' [31 (36%) control subjects versus 18 (19%) intervention subjects, $P = 0.02$]. HIV seroprevalence among the IDU subjects was 35%.

Table 1 Sex behaviors of narcology hospital patients in the Russian Partnership to Reduce the Epidemic Via Engagement in Narcology Treatment (PREVENT) randomized controlled trial at 6 months post-randomization ($n = 144$).

Sex behavior measures during the past 3 months	*Unadjusted median group differences (95% CI)*	*P-value*	*Adjusted median group differences* (95% CI)*	*P-value*
Percentage of safe sex episodes†‡	12.7 (0.0,36.6)	0.01	22.8 (−1.5,47.0)	0.07
Number of unsafe sex episodes†	−2.5 (−7.0,0.0)	0.04	−3.5 (−7.2,0.1)	0.06
	Unadjusted OR (95% CI) proportions		*Adjusted OR (95% CI)**	
Period of no unprotected sex†§	1.6 (0.8,3.1)	0.18	1.5 (0.7,3.3)	0.26
Any condom use¶	2.5 (1.1,5.5)	0.03	3.7 (1.5,8.9)	0.004

*Adjusted for baseline report of sex trade involvement and baseline value of outcome. †$n = 66$ control; $n = 73$ intervention. ‡Safe sex episode defined as condom use during anal or vaginal intercourse (yes/no). §No unprotected sex defined as condom use for 100% of sexual episodes (anal and vaginal) with all partners or abstinence for past 3 months (yes/no). ¶$n = 67$ control; $n = 77$ intervention. Median number of total sex episodes; control = 28; intervention = 20. CI: confidence interval; OR: odds ratio.

Follow-up and receipt of intervention

Follow-up was 90% (162/181) at 3 months and 80% (144/181) at 6 months, with no differential follow-up between randomization groups. Subjects lost to follow-up at the 6-month assessment were generally similar to those who were followed, except that those lost were more likely to be married (54% versus 28%, $P = 0.003$) and have a primary sex partner (95% versus 71%, $P = 0.003$). Due to loss to follow-up or subject death, 77 intervention and 67 control group subjects were included in the 6-month analyses. Of the 94 intervention subjects, 50 received the entire intervention (two in-hospital sessions and three monthly booster calls), while 44 received a partial intervention. Fifty-one, 21, 12 and 10 subjects received three, two, one and no booster sessions, respectively.

RESULTS

Primary outcomes

Percentage of safe sex

Russian PREVENT intervention subjects had a higher median percentage of safe sex episodes than control subjects at the 6-month follow-up visit (unadjusted median difference in percentage of safe sex episodes 12.7, $P = 0.01$) in unadjusted analyses (Table 1). A significant intervention effect remained even after adjusting for baseline differences in report of sex trade (adjusted median difference in percentage of safe sex episodes 29%, $P = 0.01$). Although the percentage of safe sex episodes reported at baseline did not differ significantly between groups, secondary analyses controlling for this potential confounding factor were conducted. Using this model, the treatment effect diminished and became marginally significant (median difference 22.8%, $P = 0.07$).

No unprotected sex

The intervention group had a higher odds of reporting no unprotected sex compared to controls, although the difference was not statistically significant [unadjusted odds ratio (OR) 1.6 for intervention versus controls, 95% confidence interval (CI) 0.8–3.1; $P = 0.18$]. The results were similar after adjusting for sex trade and reporting no unprotected sex (past 3 months) at baseline (adjusted OR = 1.5 for intervention versus controls, 95% CI 0.7–3.3, $P = 0.26$). Among the 48 subjects who reported no unprotected sex, nine of these (19%) were sexually abstinent (five controls, four intervention group).

Secondary outcomes

Unsafe sex episodes

At the 6-month follow-up, control subjects reported a higher median number of unsafe sex episodes in the past 3 months compared to the intervention subjects (unadjusted median difference −2.5, $P = 0.04$). However, the treatment effect became borderline significant after adjusting for baseline number of unsafe sex episodes and sex trade (adjusted median difference −3.5, $P = 0.06$).

Any condom use

The intervention subjects had a higher odds of reporting any condom use during the past 3 months (unadjusted OR 2.5 for intervention versus controls, 95% CI 1.1–5.5, $P = 0.03$). The treatment effect on any condom use persisted after adjusting for sex trade and any condom use at baseline [adjusted odds ratio (AOR) = 3.7, 95% CI: 1.5–8.9, $P < 0.01$].

No treatment differences were detected at the 3-month follow-up. However, both the intervention and

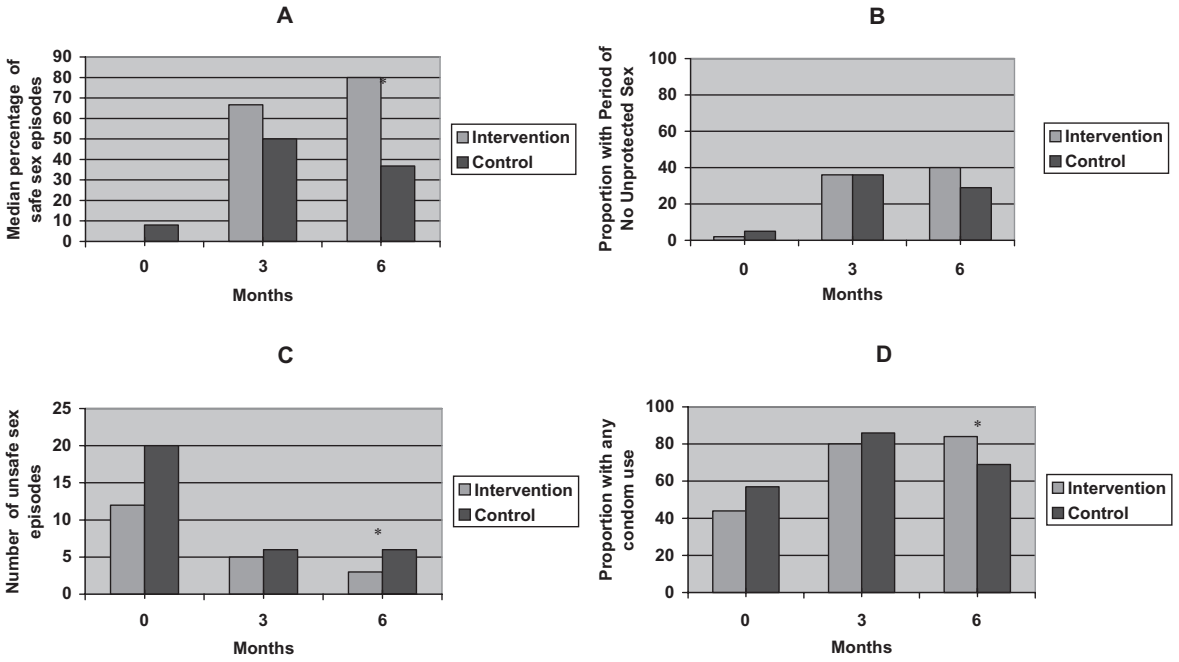


Figure 1 The effect of the Partnership to Reduce the Epidemic Via Engagement in Narcology Treatment (PREVENT) intervention on median percentage of safe sex episodes (A), percentage with periods of no unprotected sex (B), number of unsafe sex episodes (C) and percentage with any condom use (D). *$P < 0.05$, unadjusted analysis

control groups had marked improvements in the percentage of safe sex episodes, no unprotected sex and any condom use between baseline and the 3-month follow-up. While the intervention group appeared to maintain or improve their safe sex behaviors at the 6-month follow-up, the control group appeared to worsen (Fig. 1).

In exploratory subgroup analyses stratified by dependence diagnosis (alcohol versus heroin), the intervention effect appeared stronger among alcohol-dependent subjects compared to heroin-dependent subjects [median percentage of safe sex: (alcohol: 90% versus 30%, $P < 0.01$, heroin: 74% versus 59%, $P = 0.49$; proportion no unprotected sex: alcohol: 48% versus 22%, $P = 0.01$, heroin: 29% versus 32%, $P = 0.85$)]. Additional exploratory analyses stratified by depressive symptoms suggest that the effect of the intervention on percentage of safe sex and no unprotected sex was stronger among those with less depressive symptoms (data not shown). PREVENT was designed to focus upon sex risk behaviors rather than substance use; alcohol and heroin dependence were already being addressed by the narcology hospital clinicians. However, in recognition of the importance of substance use in the transmission of HIV we performed *post-hoc* exploratory analyses, which showed no significant differences between groups in injection drug use [OR = 1.2 (0.6, 2.5), $P = 0.64$] or risky alcohol consumption [0.6 (0.2, 2.1), $P = 0.38$] at the 6-month follow-up.

DISCUSSION

The results of the Russian PREVENT trial demonstrate that an HIV prevention intervention targeting sexual behaviors of alcohol and drug users is feasible in inpatient substance abuse treatment settings and suggest that it is effective in increasing any condom use. A clinically important intervention effect was observed in the hypothesized direction for the other primary outcome, a 3-month period of no unprotected sex; however, the effect was not statistically significant.

Identification of an effective sexual risk reduction program in Russia is particularly valuable, as heterosexual transmission is the next anticipated phase of the HIV epidemic driven heretofore by injection drug use [47]. A limited number of HIV behavioral interventions are documented to be effective in this region of the world; few address sexual risk, and none address sex risk in individuals with addictions [16,48]. The Russian narcology hospital yielded a cohort with risky sexual behavior, confirming the need and providing a setting for an effective sexual risk reduction intervention addressing this high-risk population.

The Russian PREVENT intervention was developed based upon an existing model demonstrated to be efficacious in US STD clinic patients [42] and recommended for dissemination by HIV prevention experts [49,50]. Few effectiveness studies have investigated whether an

adapted STD prevention model can produce desired outcomes in new settings [50]. Factors such as inadequate adherence to the originally evaluated programs or inadequate tailoring to the target population make these studies difficult to conduct [50]. Contributing to the success of our adaptation was the likelihood that the adapted model was able to remain faithful to the core elements of the original efficacious model while still being culturally and contextually appropriate.

The intervention appeared to be more effective in alcohol-dependent patients than in drug-dependent patients. This finding is surprising, given the approximate 35% HIV prevalence among IDUs in the narcology hospitals and previous research suggesting that knowledge of positive HIV serostatus reduces unsafe sex [27]. Of the 15 control participants who received brief posttest counseling due to their positive HIV infection status, 13 were IDUs; this exposure may have attenuated the difference in this relatively small subgroup analysis.

Russia's mass media campaign to encourage condom use may be valuable. However, agencies such as the World Health Organization and the US Centers for Disease Control and Prevention recommend that national HIV prevention efforts include both mass efforts (e.g. media) and intensive interventions targeted towards those at greatest risk for infection and transmission [51,52]. Targeted strategies that tailor interventions to personal HIV risk using a 'teachable moment' (i.e. a time of heightened personal risk perception, such as during an HIV or STD test) are believed to be particularly effective [51,53]. Such 'intensive' prevention interventions combined with HIV testing, such as the Russian PREVENT, may be particularly advantageous in populations who have very high HIV risk (e.g. substance-dependent people) [54,55].

Interestingly, exploratory analyses suggested that the intervention may be more effective at increasing the percentage of safe sex and no unprotected sex among those with less depressive symptoms. Future interventions should address the relationship of psychiatric comorbidities on HIV risk reduction.

When comparing the results of the current study with the RESPECT study, we observed similar magnitudes of effect for the outcome no unprotected sex; however, our study did not find a statistically significant effect on this outcome while the original study of 5758 subjects did. In RESPECT, subjects in the intervention arms were more likely to report no unprotected sex compared to the control arm at the 6-month visit [39% enhanced counseling versus 34% didactic messages; relative risk (RR) 1.14; 95% CI, 1.01–1.28; and 39% brief counseling versus 34% didactic messages; RR, 1.12; 95% CI, 1.00–1.25]. The lack of statistical significance in the current study may be an issue of statistical power.

This study's findings are consistent with US research indicating that patients in detoxification centers are at high risk for STDs, including HIV, and that sexual risk reduction programs in these settings can be efficacious [31,32,56]. A recent meta-analysis of US research demonstrated that more effective HIV prevention programs were comprised of comprehensive and fairly intensive program 'packages', including community-based outreach, substance abuse treatment, sterile syringe access and enhanced HIV/STD counseling and testing [32]. It is important to explore the utility of additional HIV prevention approaches for substance users that take into account the limited resources and existing systems available—as is the case for this study in Russia. This study provides insights not only on potential interventions for the Russian narcology hospital context; it also contributes to the growing work on the utility of brief risk reduction interventions for any patient in addiction treatment. Research up to this point has suggested some success, but has been inconclusive due to the small number of efficacy and effectiveness studies [57,58]. Notably, no previous comparable work has been conducted in eastern Europe.

At 3 months, there appeared to be an increase in safe sex in the control as well as the intervention group. Assessments conducted by telephone rather than ACASI showed improvements in sex risk behaviors in both the control and intervention groups (Fig. 1). This may have been due to factors such as exposure of the control participants to the extensive initial assessment, including an ACASI, availability of condoms (distributed to all subjects) or regression to the mean. Despite early changes, the control subjects' behavior returned toward baseline in the second 3 months, while the intervention group continued to improve. The findings of delayed sexual risk reduction effects observed in this study are consistent with previous HIV intervention research [59–62], and may perhaps be attributed to greater opportunity for intervention participants to change behavior over time [61].

The study had some major strengths: demonstration that the PREVENT intervention could be implemented in two Russian narcology hospitals supports the strong likelihood for translation of this research into practice, and ability to engage this high-risk population at a 'reachable moment' (e.g. addiction treatment) in addition to a 'teachable moment'. In St Petersburg, talented and well-trained personnel were available to provide the intervention, and similar personnel may exist elsewhere in these clinical settings. Another study strength was the heterogeneity of the research subjects in terms of gender, HIV status and substance use, supporting the notion that these results may be generalizable to the narcology patient population elsewhere.

The trial also had some limitations. Although impressive changes were reported using state-of-the-art methodology for assessing behavior change, behaviors were self-reported and participants could not be blinded to intervention status, allowing the possibility of social desirability bias. Although we were not able to obtain objective biological outcomes, we attempted to limit this bias by using ACASI technology and by using research associates who were not involved in delivering the intervention to assess outcomes. We were unable to address the number of safe sex acts with partners with discordant HIV serostatus, as we did not ask the subjects to identify the HIV serostatus of their sex partners. Also, the narcology hospital setting has a disproportionate number of men and a minority of HIV-infected patients, thus we were unable to address gender or HIV status in stratified analyses. Additionally, the international setting of this behavioral intervention study presented certain challenges to assessing the fidelity of the intervention as adapted to the Russian setting. Finally, our study was not designed to detect small-to-moderate treatment differences with high power. Moreover, adjustment for additional covariates resulted in further reduction of power. Nevertheless, all primary and secondary outcomes show clinically important differences in the hypothesized direction and the adjusted results are marginally significant for percentage of safe sex events and periods of unsafe sex, and statistically significant for any condom use.

In summary, this randomized controlled trial suggests that adaptation of a pragmatic, HIV prevention intervention may reduce risky sexual behaviors in substance-dependent patients attending Russian narcology hospitals. Dissemination of this effective intervention should be considered as a component of a broad strategy aimed at reducing HIV infections in eastern Europe and other settings.

Acknowledgements

We would like to thank the interventionists and research associates at the Russian narcology hospitals for their contributions to the study and Erika Edwards MPH, of the Data Coordinating Center at Boston University School of Public Health, for her critical reviews of the manuscript. We would also like to express our gratitude to Marina Tsoy for her work translating all study documents. These individuals received compensation from the National Institutes of Health (NIH) grant for their contributions. Grant support for this study came from the National Institute on Alcohol Abuse and Alcoholism (NIAAA), NIH: R21-AA014821 (HIV Prevention Partnership in Russian Alcohol Treatment—NCT00183118; NIH AA014821) and K24-AA015674 (Impact of Alcohol Use on HIV Infection—in United States and Russia). This study was registered with ClinicalTrials.gov through the National Institutes of Health.

Declaration of interest

The authors have no conflicts of interest to declare.

References

1. Joint United Nations Programme on HIV/AIDS (UNAIDS). *Russian Federation*. 2006. Available at: http://www.unaids.org/en/Regions_Countries/Countries/russian_federation.asp (accessed 8 March 2006).
2. World Health Organization. *Russian Federation: Summary Country Profile for HIV/AIDS Treatment Scale up*. Geneva: World Health Organization; 2005. Available at: http://www.who.int/3by5/support/june2005_rus.pdf (accessed 27 March 2006).
3. Krupitsky E., Zvartau E., Karandashova G., Horton N. J., Schoolwerth K. R., Bryant K. *et al*. The onset of HIV infection in the Leningrad region of Russia: a focus on drug and alcohol dependence. *HIV Med* 2004; **5**: 30–3.
4. Dehne K. L., Pokrovskiy V., Kobyshcha Y., Schwartlander B. Update on the epidemics of HIV and other sexually transmitted infections in the newly independent states of the former Soviet Union. *AIDS* 2000; **14**: S75–84.
5. Hamers F. F., Downs A. M. HIV in central and eastern Europe. *Lancet* 2003; **361**: 1035–44.
6. World Health Organization. *Global Status Report on Alcohol 2004*. Geneva: World Health Organization; 2004. Available at: http://whqlibdoc.who.int/publications/2004/9241562722_(425KB).pdf (accessed 16 May 2008).
7. Zohoori N., Blanchette D., Popkin B. *Monitoring Health Conditions in the Russian Federation: The Russia Longitudinal Monitoring Survey 1992–2004*. North Carolina, University of North Carolina, 2005. Available at: http://www.cpc.unc.edu/rlms/papers/health_04.pdf# (accessed 16 May 2008).
8. Avins A. L., Woods W. J., Lindan C. P., Hudes E. S., Clark W., Hulley S. B. HIV infection and risk behaviors among heterosexuals in alcohol treatment programs. *JAMA* 1994; **271**: 515–18.
9. Mahler J., Yi D., Sacks M., Dermatis H., Stebinger A., Card C. *et al*. Undetected HIV infection among patients admitted to an alcohol rehabilitation unit. *Am J Psychiatry* 1994; **151**: 439–40.
10. Bagnall G., Plant M., Warwick W. Alcohol, drugs and AIDS-related risks: results from a prospective study. *AIDS Care* 1990; **2**: 309–17.
11. Leigh B., Temple M., Trocki K. The relationship of alcohol use to sexual activity in a U.S. national sample. *Soc Sci Med* 1994; **39**: 1527–35.
12. Rees V., Saitz R., Horton N. J., Samet J. H. Association of alcohol consumption with HIV sex- and drug-risk behaviors among drug users. *J Subst Abuse Treat* 2001; **21**: 129–34.
13. Stoltz D. A., Nelson S., Kolls J. K., Zhang P., Bohm R. P., Murphey-Corb M. *et al*. Effects of *in vitro* ethanol on tumor necrosis factor-alpha production by blood obtained from simian immunodeficiency virus-infected rhesus macaques. *Alcohol Clin Exp Res* 2002; **26**: 527–34.
14. Bagby G. J., Zhang P., Purcell J. E., Didier P. J., Nelson S. Chronic binge ethanol consumption accelerates progression of simian immunodeficiency virus disease. *Alcohol Clin Exp Res* 2006; **30**: 1781–90.

15. Benotsch E. G., Somlai A. M., Pinkerton S. D., Kelly J. A., Ostrovski D., Gore-Felton C. *et al*. Drug use and sexual risk behaviours among female Russian IDUs who exchange sex for money or drugs. *Int J STD AIDS* 2004; **15**: 343–7.
16. Kozlov A. P., Shaboltas A. V., Toussova O. V., Verevochkin S. V., Masse B. R., Perdue T. *et al*. HIV incidence and factors associated with HIV acquisition among injection drug users in St Petersburg, Russia. *AIDS* 2006; **20**: 901–6.
17. Raj A., Saitz R., Cheng D. M., Winter M., Samet J. H. Associations between alcohol, heroin, and cocaine use and high risk sexual behaviors among detoxification patients. *Am J Drug Alcohol Abuse* 2007; **33**: 169–78.
18. Shaboltas A. V., Toussova O. V., Hoffman I. F., Heimer R., Verevochkin S. V., Ryder R. W. *et al*. HIV prevalence, sociodemographic, and behavioral correlates and recruitment methods among injection drug users in St. Petersburg, Russia. *J Acquir Immune Defic Syndr* 2006; **41**: 657–63.
19. Mansergh G., Purcell D. W., Stall R., McFarlane M., Semaan S., Valentine J. *et al*. CDC consultation on methamphetamine use and sexual risk behavior for HIV/STD infection: summary and suggestions. *Public Health Rep* 2006; **121**: 127–32.
20. Rhodes T., Judd A., Mikhailova L., Sarang A., Khutorskoy M., Platt L. *et al*. Injecting equipment sharing among injecting drug users in Togliatti City, Russian Federation: maximizing the protective effects of syringe distribution. *J Acquir Immune Defic Syndr* 2004; **35**: 293–300.
21. Somlai A. M., Kelly J. A., Benotsch E., Gore-Felton C., Ostrovski D., McAuliffe T. *et al*. Characteristics and predictors of HIV risk behaviors among injection-drug-using men and women in St. Petersburg, Russia. *AIDS Educ Prev* 2002; **14**: 295–305.
22. Takacs J., Amirkhanian Y. A., Kelly J. A., Kirsanova A. V., Khoursine R. A., Mocsonaki L. 'Condoms are reliable but I am not': a qualitative analysis of AIDS-related beliefs and attitudes of young heterosexual adults in Budapest, Hungary, and St. Petersburg Russia. *Cent Eur J Public Health* 2006; **14**: 59–66.
23. Merson M. H., Dayton J. M., O'Reilly K. Effectiveness of HIV prevention interventions in developing countries. *AIDS* 2000; **14**: 68–84.
24. Crepaz N., Lyles C. M., Wolitski R. J., Passin W. F., Rama S. M., Herbst J. H. *et al*. Do prevention interventions reduce HIV risk behaviours among people living with HIV? A meta-analytic review of controlled trials. *AIDS* 2006; **20**: 143–57.
25. Card J. J., Benner T., Shields J. P., Feinstein N. The HIV/AIDS Prevention Program Archive (HAPPA): a collection of promising prevention programs in a box. *AIDS Educ Prev* 2001; **13**: 1–28.
26. Noar SM. Behavioral interventions to reduce HIV-related sexual risk behavior: review and synthesis of meta-analytic evidence. *AIDS Behav* 2008; **12**: 335–53.
27. Marks G., Crepaz N., Senterfitt J. W., Janssen R. S. Meta-analysis of high-risk sexual behavior in persons aware and unaware they are infected with HIV in the United States: implications for HIV prevention programs. *J Acquir Immune Defic Syndr* 2005; **39**: 446–53.
28. Weinhardt L. S., Carey M. P., Johnson B. T., Bickham N. L. Effects of HIV counseling and testing on sexual risk behavior: a meta-analytic review of published research, 1985–1997. *Am J Public Health* 1999; **89**: 1397–405.
29. D'Aunno T., Vaughn T. E., McElroy P. An institutional analysis of HIV prevention efforts by the nation's outpatient drug abuse treatment units. *J Health Soc Behav* 1999; **40**: 175–92.
30. Gibson D. R., McCusker J., Chesney M. Effectiveness of psychosocial interventions in preventing HIV risk behaviour in injecting drug users. *AIDS* 1998; **12**: 919–29.
31. Semaan S., Des Jarlais D. C., Sogolow E., Johnson W. D., Hedges L. V., Ramirez G. *et al*. A meta-analysis of the effect of HIV prevention interventions on the sex behaviors of drug users in the United States. *J Acquir Immune Defic Syndr* 2002; **30**: S73–S93.
32. Prendergast M. L., Urada D., Podus D. Meta-analysis of HIV risk-reduction interventions within drug abuse treatment programs. *J Consult Clin Psychol* 2001; **69**: 389–405.
33. Burrows D., Sarankov I. [A decrease in harm: a new concept for Russian public health]. *Zh Mikrobiol Epidemiol Immunobiol* 1999; **1**: 107–8.
34. Burrows D. Strategies for dealing with HIV/AIDS in the former Soviet Union. *Dev Bull* 2000; **52**: 49–51.
35. Open Society Institute. *Drugs, AIDS, and Harm Reduction: How to Slow the HIV Epidemic in Eastern Europe and Former Soviet Union*. 2001. Available at: http://www.soros.org/initiatives/health/focus/ihrd/articles_publications/publications/drugsaidshr_20010101 (accessed 24 October 2006).
36. Metzger D., Woody G. E., McLellan A., O'Brien C., Druley C., Navaline H. *et al*. Human immunodeficiency virus seroconversion among intravenous drug users in and out of treatment: an 18-month prospective follow-up. *J Acquir Immune Defic Syndr* 1993; **6**: 1049–56.
37. Sullivan L. E., Metzger D. S., Fudala P. J., Fiellin DA. Decreasing international HIV transmission: the role of expanding access to opioid agonist therapies for injection drug users. *Addiction* 2005; **100**: 150–8.
38. Ksobiech K., Somlai A. M., Kelly J. A., Gore-Felton C., Benotsch E., McAuliffe T. *et al*. Demographic characteristics, treatment history, drug risk behaviors, and condom use attitudes for U.S. and Russian injection drug users: the need for targeted sexual risk behavior interventions. *AIDS Behav* 2005; **9**: 111–20.
39. Moher D., Schulz K. F., Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA* 2001; **285**: 1987–91.
40. Radloff L. The CES-D scale: a self-report depression scale for research in the general population. *Appl Psychol Meas* 1977; **1**: 401.
41. Ware J., Sherbourne C. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992; **30**: 473–83.
42. Kamb M. L., Fishbein M., Douglas J. M. Jr, Rhodes F., Rogers J., Bolan G. *et al*. Efficacy of risk-reduction counseling to prevent human immunodeficiency virus and sexually transmitted diseases: a randomized controlled trial. Project RESPECT Study Group. *JAMA* 1998; **280**: 1161–7.
43. Navaline H. A., Snider E. C., Petro C. J., Tobin D., Metzger D., Alterman A. I. *et al*. Preparations for AIDS vaccine trials. An automated version of the Risk Assessment Battery (RAB): enhancing the assessment of risk behaviors. *AIDS Res Hum Retroviruses* 1994; **10**: 281–3.
44. Weinhardt L. S., Carey M. P., Maisto S. A., Carey K. B., Cohen M. M., Wickramasinghe S. M. Reliability of the timeline follow-back sexual behavior interview. *Ann Behav Med* 1998; **20**: 25–30.

45. Sobell L., Sobell M. *Alcohol Timeline Followback (TLFB) Users' Manual*. Toronto, Canada: Addiction Research Foundation; 1995.
46. McLellan A., Luborsky L., Woody G., O'Brien C. P. An improved diagnostic evaluation instrument for substance abuse patients. The Addiction Severity Index. *J Nerv Ment Dis* 1980; **168**: 26–33.
47. Karon J. M., Fleming P. L., Steketee R. W., De Cock K. M. HIV in the United States at the turn of the century: an epidemic in transition. *Am J Public Health* 2001; **91**: 1060–8.
48. Luo R. F., Cofrancesco J. Jr. Injection drug use and HIV transmission in Russia. *AIDS* 2006; **20**: 935–6.
49. National Institutes of Health. *National Institutes of Health Fiscal Year 2006 Plan for HIV-related Research*. 2004. Available at: http://www.oar.nih.gov/public/pubs/fy2006/00_Overview_FY2006.pdf (accessed 24 October 2006).
50. Solomon J., Card J. J., Malow RM. Adapting efficacious interventions: advancing translational research in HIV prevention. *Eval Health Prof* 2006; **29**: 162–94.
51. Sumartojo E., Carey J. W., Doll L. S., Gayle H. Targeted and general population interventions for HIV prevention: towards a comprehensive approach. *AIDS* 1997; **11**: 1201–9.
52. World Health Organization. *Policy and Programming Guide for HIV/AIDS Prevention and Care among Drug Users*. Geneva, Switzerland: WHO Press; 2005. Available at: http://wwwint/hiv/pub/idu/iduguide/en/ (accessed 24 October 2006).
53. Myers T., Worthington C., Haubrich D. J., Ryder K., Calzavara L. HIV testing and counseling: test providers' experiences of best practices. *AIDS Educ Prev* 2003; **15**: 309–19.
54. Joint United Nations Programme on HIV/AIDS (UNAIDS), World Health Organization (WHO). *UNAIDS/WHO Policy Statement on HIV Testing*. 2004. Available at: http://wwwint/hiv/pub/vct/statement/en/ (accessed 24 October 2006).
55. World Health Organization, Joint United Nations Programme on HIV/AIDS (UNAIDS). *Advocacy Guide: HIV/AIDS Prevention among Injection Drug Users*. Geneva, Switzerland: World Health Organization; 2004. Available at: http://www.who.int/hiv/pub/idu/iduadvocacyguide/en/index.html (accessed 24 October 2006).
56. Avins A. L., Lindan C. P., Woods W. J., Hudes E. S., Boscarino J. A., Kay J. *et al.* Changes in HIV-related behaviors among heterosexual alcoholics following addiction treatment. *Drug Alcohol Depend* 1997; **44**: 47–55.
57. Stein M. D., Anderson B., Charuvastra A., Maksad J., Friedmann P. D. A brief intervention for hazardous drinkers in a needle exchange program. *J Subst Abuse Treat* 2002; **22**: 23–31.
58. Dunn C., Deroo L., Rivara F. P. The use of brief interventions adapted from motivational interviewing across behavioral domains: a systematic review. *Addiction* 2001; **96**: 1725–42.
59. Jemmott J. B. III, Jemmott L. S., Braverman P. K., Fong G. T. HIV/STD risk reduction interventions for African American and Latino adolescent girls at an adolescent medicine clinic: a randomized controlled trial. *Arch Pediatr Adolesc Med* 2005; **159**: 440–9.
60. Kalichman S. C., Rompa D., Cage M., DiFonzo K., Simpson D., Austin J. *et al.* Effectiveness of an intervention to reduce HIV transmission risks in HIV-positive people. *Am J Prev Med* 2001; **21**: 84–92.
61. Raj A., Amaro H., Cranston K., Martin B., Cabral H., Navarro A. *et al.* Is a general women's health promotion program as effective as an HIV-intensive prevention program in reducing HIV risk among Hispanic women? *Public Health Rep* 2001; **116**: 599–607.
62. Villarruel A. M., Jemmott J. B. III, Jemmott L. S. A randomized controlled trial testing an HIV prevention intervention for Latino youth. *Arch Pediatr Adolesc Med* 2006; **160**: 772–7.

Addiction

Commentary

CAN ONE BE AN EXPERT IN ADDICTION MEDICINE WITHOUT EXPERTISE IN PAIN MANAGEMENT?

The paper by Caldeiro *et al.* [1] in this issue notes the conundrum in which patients with pain and addiction find themselves: 'pain complaints are difficult to address until addiction is treated, but addiction treatment is compromised by persistent pain'. Patients share this conundrum with their health-care providers. We see this Catch 22 situation in the hospital, medical and mental health clinics and addiction treatment programs, yet the prevalence and outcomes of these co-occurring, biologically linked conditions has been described relatively recently in the medical literature. Furthermore, our clinical management is lacking in evidence-based guidance.

Caldeiro *et al.* demonstrate, among 582 patients with non-opioid substance dependence in the Veterans Administration health-care system, that pain is common and it matters. Patients with persistent pain have worse medical and substance use outcomes and use more resources. This study confirms and extends earlier work [2,3] by demonstrating this relationship in addictions other than opioid dependence. A previous study among patients in methadone maintenance and short-term residential treatment programs found the prevalence of chronic severe pain to be 37% and 24%, respectively, and use of illicit drugs to treat their pain similarly common (34% and 51%, respectively).

Characterizing the overlap between pain and addiction beyond opioids is important, as it is now clear that the relationship of pain and addiction is more complicated than one family of opioid receptors (e.g. mu). For example, pre-clinical research demonstrates that there is substantial overlap among the opioid and cannabinoid receptor systems in modulating analgesia both in inflammatory and neuropathic pain conditions [4]. Clinical trials of cannabinoid agonists show promise for chronic neuropathic pain [5,6]. Alcohol is a weak analgesic. Cocaine, while not a known analgesic in humans, was first used clinically as a topical anesthetic in the 1880s. The abuse of both has been associated with prescription opioid abuse among patients with chronic pain [7,8]. Do these associations between pain and different addictions represent a single phenomenon mediated by a final common neurochemical pathway, or are they driven separately by each substance, pain source and individual's vulnerability? Is it best to couple known treatments of addiction and pain, or are there better therapies targeted specifically to patients with both conditions?

Future research will address these questions, providing a more complete understanding of the interplay between pain and addiction. The treatment of chronic pain with opioids has a modest evidence base [9–11]. Most studies have follow-up periods of less than 6 months, involve highly selected populations and do not include patients with co-occurring addiction. Intervention studies targeted to individuals with both addictions and chronic pain are rare. The management recommendations in the *Principles of Addiction Medicine* [12] repeat the approach and clinical guidelines developed in the early 1990s for the management of cancer pain by the Agency for Health Care Policy Research. These guidelines have not been examined specifically in patients with underlying addictions with chronic non-cancer pain. The evidence we have to address these interwoven conditions is based largely on studies focused upon these conditions in isolation. It is time for rigorously designed longitudinal research of behavioral and pharmacological treatments focused upon groups with addictions and overlapping chronic pain.

Despite the ever-present need for future research, the implications of these authors' findings speak to the need for present action. Pain assessments and efforts to address painful conditions adequately should be explored, as noted by Caldeiro *et al.* in their discussion. Models of care that best achieve such assessment and appropriate relief will improve current clinical treatment in both substance use and medical settings.

Beyond improved diagnostic tools and treatment models, these findings call for physicians who practice addiction medicine to become expert in the evaluation and treatment of pain. The fact that pain is present in more than half our patients with addiction, even in those with non-opioid addictions, begs the question: can one be an expert in addiction medicine without expertise in pain management? Despite the current modest evidence base, evaluation and treatment of pain in addicted patients merits a place among the core competencies in addiction medicine. The core competencies for a subspecialty in addiction psychiatry call for addiction psychiatrists to demonstrate knowledge of the evaluation and consultation on chronic pain [13]. As it considers the content of a certification examination, the new American Board of Addiction Medicine (ABAM) should include pain assessment and management as one of its core subjects. Such a mandate will support and propel the drive for stronger evidence to treat our addicted patients with painful conditions.

Declarations of interest

None.

Keywords Addiction, certification examination, medical education, non-opioids, opioids, pain.

JEFFREY H. SAMET & ALEXANDER Y. WALLEY
Boston University School of Medicine, Boston, MA, USA.
E-mail: jsamet@bu.edu

References

1. Caldeiro R. M., Malte C. A., Calsyn D. A., Baer J. S., Nichol P., Kivlahan D. R. *et al.* The association of persistent pain with out-patient addiction treatment outcomes and service utilization. *Addiction* 2008; **103**: 1996–2005.
2. Larson M. J., Paasche-Orlow M., Cheng D. M., Lloyd-Travaglini C., Saitz R., Samet J. H. Persistent pain is associated with substance use after detoxification: a prospective cohort analysis. *Addiction* 2007; **102**: 752–60.
3. Rosenblum A., Joseph H., Fong C., Kipnis S., Cleland C., Portenoy R. K. Prevalence and characteristics of chronic pain among chemically dependent patients in methadone maintenance and residential treatment facilities. *JAMA* 2003; **289**: 2370–8.
4. Manzanares J., Julian M., Carrascosa A. Role of the cannabinoid system in pain control and therapeutic implications for the management of acute and chronic pain episodes. *Curr Neuropharmacol* 2006; **4**: 239–57.
5. Karst M., Salim K., Burstein S., Conrad I., Hoy L., Schneider U. Analgesic effect of the synthetic cannabinoid CT-3 on chronic neuropathic pain: a randomized controlled trial. *JAMA* 2003; **290**: 1757–62.
6. Rog D. J., Nurmikko T. J., Friede T., Young C. A. Randomized, controlled trial of cannabis-based medicine in central pain in multiple sclerosis. *Neurology* 2005; **65**: 812–19.
7. Ives T. J., Chelminski P. R., Hammett-Stabler C. A., Malone R. M., Perhac J. S., Potisek N. M. *et al.* Predictors of opioid misuse in patients with chronic pain: a prospective cohort study. *BMC Health Serv Res* 2006; **6**: 46. Available from: http://www.biomedcentral.com/1472-6963/6/46 (accessed 20 October 2008).
8. Turk D. C., Swanson K. S., Gatchel R. J. Predicting opioid misuse by chronic pain patients: a systematic review and literature synthesis. *Clin J Pain* 2008; **24**: 497–508.
9. Trescot A. M., Helm S., Hansen H., Benyamin R., Glaser S. E., Adlaka R. *et al.* Opioids in the management of chronic non-cancer pain: an update of American Society of the Interventional Pain Physicians' (ASIPP) Guidelines. *Pain Physician* 2008; **11**: S5–62.
10. Trescot A. M., Glaser S. E., Hansen H., Benyamin R., Patel S., Manchikanti L. Effectiveness of opioids in the treatment of chronic non-cancer pain. *Pain Physician* 2008; **11**: S181–200.
11. Deshpande A., Furlan A., Mailis-Gagnon A., Atlas S., Turk D. Opioids for chronic low-back pain. *Cochrane Database Syst Rev* 2007; Issue 3. Art. No.: CD004959. DOI: 10.1002/14651858.CD004959.pub3.
12. Savage S. R. Principles of Pain Management in the Addicted Patient. In: Graham A. W., Schultz T. K., Mayo-Smith M. F., Ries R. K., Wilford B. B., editors. *Principles of Addiction Medicine*, 3rd edn. Chevy Chase, MD: American Society of Addiction Medicine; 2003, p. 1405–19.
13. American Board of Psychiatry and Neurology. *The American Board of Psychiatry and Neurology Addiction Psychiatry Core Competencies Outline Version 3*. Available at: http://www.abpn.com/competencies.htm (accessed 3 September 2008).

JOURNAL OF WOMEN'S HEALTH
Volume 17, Number 2, 2008

DOI: 10.1089/jwh.2007.0083

Healthcare Preferences Among Lesbians: A Focus Group Analysis

MARGARET R. SEAVER, M.D., M.P.H.,[1,2,3] KAREN M. FREUND, M.D., M.P.H.,[2,3] LESLIE M. WRIGHT, M.P.H.,[2,3] JENNIFER TJIA, M.D., M.S.C.E.,[4] and SUSAN M. FRAYNE, M.D., M.P.H.[5,6,7]

ABSTRACT

Objective: **The healthcare needs of lesbians are not well understood. We sought to characterize lesbians' experiences with, and preferences for, women's healthcare.**

Methods: **We conducted three age-stratified focus groups (18–29, 30–50, and >50 years) with a total of 22 participants using a semistructured interview guide to elicit lesbians' experiences and preferences. We analyzed transcripts of these audiotaped sessions using the constant comparative method of grounded theory. Community-dwelling women who self-identified as lesbian and responded to advertisements were selected on first-come basis.**

Results: **Participants voiced experiences and preferences for healthcare that emerged into three themes: desired models of care, desired processes of care, and desired patient-provider relationship. Each theme was further developed into multiple subthemes. Within the subthemes we identified issues that were specific to lesbians and those that were general women's health issues. Participants preferred, but did not always receive, care that is comprehensive in scope, person centered, nondiscriminatory, and inclusive of them as lesbians.**

Conclusions: **Healthcare providers, institutions, and society should adopt an inviting, person-centered approach toward lesbians seeking healthcare, assure them access to healthcare information, and establish healthcare delivery systems that take all aspects of health into account.**

INTRODUCTION

LESBIANS ARE AN IMPORTANT and underrecognized patient population about which little is known. The Institute of Medicine (IOM) has specifically addressed the need for research in this population by stating that studies are needed "[to] identify possible barriers to mental and physical healthcare services [for lesbians] and ways to increase their access to these services."[1(p10)] Without

[1]VA Boston Healthcare System, Boston, Massachusetts.
[2]National Center of Excellence in Women's Health, Boston University School of Medicine, Boston, Massachusetts.
[3]Women's Health Unit, Section of General Internal Medicine, Department of Medicine, Boston University Medical Center, Boston, Massachusetts.
[4]Division of Geriatric Medicine, Department of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania.
[5]Center for Health Care Evaluation, VA Palo Alto Health Care System, Palo Alto, California.
[6]Division of General Internal Medicine, and [7]Center for Primary Care and Outcomes Research, Stanford University School of Medicine, Palo Alto, California.

Support for this research was provided by contracts from the U.S. Public Health Service, Office of Women's Health, and by the U.S. Department of Veterans Affairs and Health Services Research and Development Service grant RCD 98-312.

The views expressed in this paper are those of the authors and do not necessarily represent the views of the Department of Veterans Affairs.

data describing the healthcare needs of lesbians, clinicians will be unable to design interventions to improve the quality of care for this group of patients who comprise 3%–10% of all U.S. women.[1–3]

Unfortunately, the literature base identifying optimal approaches to the delivery of lesbian health care is thin. The available literature supports the assertion that lesbians use healthcare less than heterosexual women do and that they may experience lower health status because of this.[1,4–7] For example, studies have documented lower rates of breast and cervical cancer screening among lesbians.[1,8] Other studies have focused on health behaviors, such as smoking and alcohol consumption, and have found that lesbians use these substances, which are closely linked to poor health outcomes, more than heterosexual women do.[7,9–11] Although some authors have pointed to access problems (e.g., lack of insurance) as key barriers to healthcare,[4,6,7] others have concluded that negative experiences with the healthcare system promote avoidance of care.[9,12,13] To inform interventions designed to address such disparities, more information about the healthcare preferences of lesbians is needed.

Prior studies on lesbian health have been limited by several methodological challenges. First, it is difficult to define who is a lesbian, and definitions differ between studies. The term "lesbian" can refer to sexual behavior (i.e., women who have sex with women, although not necessarily exclusively with women), sexual orientation (i.e., women who have a sexual interest in other women, regardless of whether they have sex with other women), and cultural identity (i.e., women who identify with a community of like-minded women).[14] This variation may lead to potential problems in generalizability and precision of cohort definitions. Second, lesbians are a heterogeneous group belonging to every age group, ethnic group, and socioeconomic class.[1] Demographic cohort selection effects may influence assessments of healthcare needs. Third, it can be difficult to identify and recruit potential research participants. Some early lesbian health studies recruited women from bars and nightclubs, limiting the generalizability of findings to women who were often smokers and alcohol users.[7] Often marginalized and subjected to social stigma,[15] lesbians recruited in other settings may hesitate to disclose their sexual orientation to investigators.[12,16] Healthcare settings, which are usually an excellent site for research participant recruitment, may be less helpful for the recruitment of lesbians because healthcare providers typically are unaware of their patients' sexual orientations.[17] Thus, for various reasons, subgroups of lesbians may not enroll in studies recruiting from healthcare settings, potentially introducing bias.

To address the call for a better understanding of lesbians' healthcare needs and attempt to address these methodological difficulties, we recruited self-identified lesbians from the community and conducted an exploratory study using focus groups. The specific aim of this study was to characterize lesbians' experiences with and expectations of women's healthcare. To our knowledge, the structure of healthcare preferred by lesbians has not been examined previously.

MATERIALS AND METHODS

Participants and recruitment

Women were eligible to participate in this study if they were aged ≥18, community dwelling, and self-identified as lesbians. Potential participants were recruited using advertisements in newspapers serving the gay and lesbian community and fliers posted at Boston area businesses and primary care practices. Three age-stratified groups (18–29, 30–50, >50 years) of up to 9 participants each were enrolled on a first-come basis. Participants were age stratified so that discussion would benefit from both their shared experience of time and world events and their shared age-based healthcare needs.[18,19] Participants received an honorarium. The study was approved by the Institutional Review Board at the Boston University School of Medicine.

Instruments

The 27-item, semistructured focus group protocol used in this study was developed jointly by six sites of the Department of Health and Human Services (DHHS) National Centers of Excellence in Women's Health.[20] Drawing on a literature review and expert input, questions on the protocol were designed to elicit knowledge, attitudes, and beliefs about women's health, the healthcare environment, and the participants' own experience with the healthcare system. The focus group for-

mat, which uses an inductive approach, has been advocated for the study of topics like this that have received little prior investigation.[21] Participants also completed a brief demographic questionnaire.

Focus groups

In separate arms of the larger National Centers of Excellence in Women's Health study, other sites applied the same protocol to different study populations.[22] We report the results of the study arm that enrolled lesbians. The focus groups were conducted in a private conference room in a major medical center by a single interviewer with expertise in qualitative research methods. The interviewer facilitated discussion, assured that all participants had an opportunity to participate, and asked for clarification or elaboration if needed but did not direct the content of participants' comments. One or two other investigators observed each focus group. Each 2.5-hour session was audiotaped and then transcribed verbatim.

Analysis

After checking of transcriptions for accuracy of content, the transcripts were analyzed using the constant comparative method of grounded theory.[23,24] In this method, the analytical process undergoes continuous refinement, repeatedly feeding back into the process of coding and iteratively honing the identified themes. Five independent reviewers analyzed transcripts in four stages. In the first stage, the five reviewers read each transcript identifying key words and phrases that represented an idea or concept expressed by the participants. Key words and phrases were in the participants' own words whenever possible. In the second stage, reviewers met to discuss their initial key words and phrase codes. When key words and phrases demonstrated a clear relationship to each other, reviewers aggregated these related ideas into themes, such as patient-provider relationship, and subthemes within each theme. In the third stage, reviewers recoded the transcripts, applying the newly developed themes and subthemes to each transcript. At this stage, transcripts were further coded according to whether the expressed ideas were specific to lesbians or applied to women's health in general. In the fourth and final stage of analysis, reviewers met to identify and review differences in coding. Differences were reconciled through consensus. Throughout this process, the reviewers repeatedly returned to the original transcripts to assure that our analysis remained grounded in the participants' ideas. In this way indigenous themes that characterized the experience of the participants were identified.[25] Analysis was facilitated by the use of the qualitative software package NUD*IST V4.0 (Qualitative Solutions and Research, Melbourne, Australia).

RESULTS

Twenty-two women who self-identified as lesbians participated (groups of 8, 9, and 5 in the youngest, middle, and oldest age groups, respectively). The mean age was 38 (range 22–63). Nine were from racial/ethnic minority groups. Twenty reported having had a healthcare visit within the past year, and 10 reported having made more than four healthcare visits in the past year. Five had no regular healthcare provider, 4 had no health insurance, 4 had less than a college degree, and 9 had household income <$35,000 (2 did not provide income information).

We identified three major themes: model of care, process of care, and the patient-provider relationship, and several subthemes within each theme. We describe each theme and subtheme, providing representative quotations to illustrate these themes and dissenting views when present.[26]

Our analysis also identified several key preferences that were specific to lesbians (Table 1). Within the theme of model of care, women preferred comprehensive care that was inclusive of lesbians, provided by clinicians who had specific knowledge about lesbians and who were able to provide health information specifically about lesbians. Within the theme of process of care, women's preferences included sufficient time to disclose information about their sexuality and assurance that their sexual orientation would be treated confidentially. Participants' preference was for settings where there were lesbian-specific materials in the waiting room and sexual orientation options on the intake forms. Within the theme of patient-provider relationship, the preference was for a patient-provider relationship that encouraged disclosure of sexual orientation (promoted by providers who are nonheterosexist

Table 1. Themes and Key Issues Specific to Lesbians

Themes and subthemes	*Sample key issues specific to lesbians*
Model of care	
Comprehensive care	Inclusive of lesbians
Incorporation of integrative health	Specific health information for lesbians
Access to health information	Specific training for providers about lesbians
Women's health training for providers	
Process of care	
Time constraints	Sufficient time with providers to disclose and discuss sexual orientation
Payment systems	
Office systems	Confidentiality especially about sexual orientation
	Lesbian-specific information in waiting rooms
	Sexual orientation options on intake forms
Patient-provider relationship	
Communication style	Open communication to enhance opportunities for disclosure and discussion of sexual orientation
Provider characteristics	Nonheterosexist, nonjudgmental providers
Patients being known as individuals	Providers attuned to lesbian-specific concerns of their patients

and nonjudgmental) and for a provider attuned to the specific concerns of lesbians.

Model of care

Model of care was one of three major themes that emerged from the analysis. Comments were categorized as model of care if they related to an overarching philosophy of healthcare delivery and the scope of care. Four subthemes were identified: (1) comprehensive care, emphasizing preventive care and mental health, (2) incorporation of integrative health into routine care, (3) access to health information, and (4) special women's health training for providers.

Comprehensive care. Some participants identified the ideal model of care as comprehensive care. This was defined as care that addressed the whole person in an integrated fashion and that was not limited to gynecological or reproductive care. This applied to acute care, preventive care, nutrition, and mental healthcare: "Sometimes, in mainstream society, when people think 'women's health' that necessarily means reproductive health and not comprehensive healthcare, one-stop shopping." Mental healthcare was identified as an important aspect of a comprehensive model of care: "My healthcare has kind of been a little separated. It's like there is mental health and then there's 'real' health. That's the way it's been presented."

Participants also identified barriers to delivering this model of women's healthcare. One idea that emerged was that the traditional definition of women's health focused on reproductive health, a paradigm that hinders the acceptance of a more comprehensive model. Another obstacle described is the use of separate clinics to deliver different types of care (e.g., urgent, preventive, reproductive, mental healthcare, as well as chronic disease management): "You go to a primary care physician, and then, do you go to a separate gynecologist? Then say you have heart trouble, do you go to a separate cardiologist? Yes, you do all those things."

Participants believed that women's health, as it is usually practiced without this comprehensive model, was not inclusive of lesbians. Most had felt marginalized by a healthcare system that emphasized reproductive health and implicitly assumed that all patients are heterosexual: "Every time I went to a gynecologist—that was about the only time I went to a doctor—I'd be the only person in the room who wasn't pregnant, and I felt I'm going to be the last person waited on here because they're never going to make any money on me; I'm not going to have any children."

The groups expressed the idea that lesbians felt excluded from healthcare. The emphasis on reproductive health gave the impression that lesbians do not need gynecological examinations. The marginalization they experienced extended to their families. Participants stated that partners were often excluded from the healthcare process in instances where a heterosexual partner would have been included. One participant had had her partner included in her healthcare: "They let my partner come in with me. They treated her nor-

mally. It was such a joy to think they had taken the time to think that I wasn't widowed or divorced and that I did have a partner and that I could include this person in my follow-up meetings."

Incorporating Integrative Health. The second subtheme expressed by the participants was the importance of including integrative health into traditional healthcare. Most participants used some form of integrative therapy, including herbal medicine, acupuncture, vitamins, massage, meditation, and chiropractors. They wanted their mainstream practitioners (doctors and nurses) to know about these practices and incorporate them into treatment plans. The general concept of integrative health appeared to be a fundamental component of most participants' conceptualizations of healthcare: "I really want it to be a place that's very integrated, where I can feel like I can bring my whole self, I can bring my questions about my yeast infections, I can bring my questions about this Chinese drug." Other participants echoed this sentiment: "And you don't have to worry what the doctor at the acupuncture center is telling you, or the herbal therapist is telling you, and then what the Western doctor is saying, that there would be integration of those aspects as well."

However, participants understood the traditional model of care as exclusive of integrative therapies: "When I first started seeing a chiropractor, I was seeing a doctor who had no use for chiropractors at all, and they would get into this back and forth, debasing each other, and telling me it was stupid to see the other person." In addition to perceived conflict between providers, a lack of insurance coverage and lack of referral sources were identified as barriers to a comprehensive model of care.

Access to Information. The third major subtheme defining model of care was access to information about health, obtained from books, the Internet, and providers. Some women (especially in the older group) commented that information about sexual development and sexual identity was difficult to obtain from any source. This served as a barrier to ideal care: "It's her body, but don't tell her about it. You had to find out [about sexual development] on the street, from your friends." Participants identified the need for more explanation of diagnoses, treatment, and procedures, especially gynecological and surgical procedures, from healthcare providers: "There were two gynecologists; I wasn't feeling that they were giving me information."

One participant stated that fliers in waiting rooms could provide the opportunity for further discussion of health issues with practitioners. Participants thought that lay sources were also good sources of information. Women identified support and informational groups as currently underused, suggesting that groups could enhance access to information on such topics as stress, menopause, and reproductive health.

Participants identified language as a barrier to exchange of information with providers. Misinformation—from the Internet, family, friends, "the street," and even unknowledgeable providers—was another barrier to good healthcare: "There's so much terrible stuff on the web. It can be a source of extremely bad, wrong, kind of information."

Participants stated that health information specific to lesbians was particularly difficult to access. Waiting rooms often lack lesbian-specific information. Providers often lack information about lesbian health, such as their risks of HIV or other sexually transmitted diseases (STDs), and their risk of breast and gynecological cancers: "I'm with a partner, and have been for two years, who's HIV positive. I'm always going to have questions. I'm concerned about different types of sex." "I've never seen any mention of a dental dam or rubber gloves in a doctor's office."

However, some women did experience women's health in specific settings as inclusive and supportive of them as lesbians. These settings included clinics devoted to the care of gay men and lesbians and offices where healthcare providers identified themselves as lesbian or gay, or advertised (by word of mouth) that they were oriented to the care of lesbians. Women perceived providers who ask about sexual orientation as more open to lesbians.

Specialized knowledge of women's health. The fourth major subtheme in the models of care theme, closely related to the need for access to information, was the importance of providers having specialized knowledge in women's health and lesbian health. Participants identified lack of specialized training as a barrier to healthcare for women in general and especially lesbians and recommended that medical school curricula fo-

cus on lesbian health, lesbian lifestyle or culture, and integrated healthcare approaches: "How many providers know about safe lesbian sex?" "Women who haven't had children, are they more or less likely to have breast cancer? There's a different constellation of health issues that are part of the lesbian community."

Process of care

Process of care was the second major theme that emerged from the analysis. Comments were categorized as process of care if they related to systems issues in the delivery of healthcare. Three major subthemes were (1) time constraints as they impacted both the patient and the provider, (2) payment systems including insurance and managed care, and (3) office systems including confidentiality, comfort in the office, and referral.

Time constraints. The first subtheme, time constraints, included limited time for appointments that impeded getting/giving a complete history and adversely affected quality of care. Providers strapped for time might not review a patient's history prior to a visit, leading to unnecessary testing: "Every time I've had a Pap smear, they have called on the phone and told me to come back and take it again, that it was no good. I tell them that I've had a hysterectomy, and they say 'Oh you did? Well, then okay.' But it scares you, when that call comes in."

Women reported that ideal care would leave them feeling like people as opposed to numbers. Insufficient time to review their history and social issues with providers led women to the feel that they were widgets on an assembly line. Women also believed that their time was as valuable as the practitioner's, and, therefore, they did not like to wait or reschedule. They equated respect for their time with respect for them as people.

Payment systems. The second major subtheme in process of care related to payment systems, including insurance issues and managed care. Most participants appreciated getting their care at a single location, both for convenience and to reduce the stress of visits to multiple providers. They wanted more information about navigating the increasingly complex medical system. They also wanted their providers to communicate with each other.

Participants believed that healthcare was difficult to access especially if they had no insurance and that this led to inappropriate emergency room use or delaying and avoiding healthcare. Incomplete coverage for comprehensive care was an additional concern. Participants wondered why integrative care, such as acupuncture and massage, was not covered by insurance.

Office systems. The third major subtheme was office systems. Participants stressed the importance of confidentiality of the medical record, confidential treatment of sensitive information by the staff (including sexual orientation), and private space to fill out history forms. They did not want their chief complaint written where it would be visible to anyone. They did not like being asked the same personal questions by multiple people, especially if the person's role in the patient's care was unclear. Participants elected not to disclose personal information, including sexual orientation, if they thought the chart was not confidential.

Another important issue was examination room comfort. Long waits, small examination gowns, gowns made of paper, and cold rooms make visits uncomfortable. Environmental factors clearly impacted women's healthcare experiences: "I don't like the time when you're in the exam room undressed, sitting there in this doily and waiting. You're looking at the stirrups; it's not a friendly sight at all. And it's usually cold."

Patient-provider relationship

The final major theme emerging from the analysis related to the patient-provider relationship. Comments were categorized as patient-provider relationship if they addressed attributes of the provider or characteristics of the provider's interactions with the patient. The three major subthemes identified were (1) communication style, (2) provider characteristics, including gender, sexual orientation, and knowledge, and (3) patients being known as individuals.

Communication style. The first subtheme was communication style. Participants highly valued solid communication skills, nonjudgmental attitudes, and nonhierarchical relationships. Women sought a provider who was willing to address difficult issues with full disclosure. Attentive listening was essential, as was accessible language: "I feel like they always want to fix things with med-

icine first. And I would like someone who is willing to talk to me and figure out what my lifestyle is, and if other things can work first before I have to take pills."

Trust was a major component of ideal care. A major barrier to a trusting patient-provider relationship stemmed from the shaming behaviors of providers, which undermined good communication. Some women dreaded healthcare visits because of the focus on poor health habits. There was the sense that providers were constantly chastising patients about weight loss, exercise, smoking cessation, and drug or alcohol use, without much of an understanding of the underlying issues in an individual patient's life that contribute to these health behaviors: "'You should eat less and exercise more and lose weight.' If it was that easy everyone would do it, it's shaming. It's the same thing with the gowns, I'm not going to say 'This gown is too small.' Those things are hard, I just don't sit around my house eating Snickers, there is a host of other things going on."

In fact, some women delayed or avoided healthcare because of poor communication: "Ideally [healthcare] would be about me in my life, and not me in my life as the doctor perceives it. My doctor is not going to be my best friend, but awareness would be something huge. I think that's one of the reasons I don't seek healthcare now."

Participants were often reluctant to disclose their sexuality to providers because of the negative reaction they received or because the participants perceived that healthcare providers might consider homosexuality a mental illness: "When you verify, you say, 'I have sex with women,' they're like, 'Oh, okay, well, we're moving on.' I think it's true. I think they just want to move as quickly as possible away from the crazy lesbian." "There are a lot of doctors who still believe that being gay is a mental illness."

Most women had not been asked about sexual orientation, and they had experienced an assumption of heterosexuality from providers. This usually occurred in the context of reproductive health. Providers either repeatedly asked questions about the need for birth control or were concerned about pregnancy despite the patient's disclosure of sexual orientation. Providers' homophobia led some to drop out of treatment: "I wanted to see how they would react when I said I was a lesbian, and if I didn't get good vibes, I was out of there."

Communication skills, attitude, and knowledge in an ideal relationship were linked to specific issues facing lesbians. For example, the way in which a lesbian was asked (or not asked) about her sexual orientation strongly impacted the patient-provider relationship. Providers who had good communication skills and were nonjudgmental and accepting were favored, in part because their attitude invited disclosure of sexual orientation. Participants believed that if a provider asked about sexual orientation either directly or via questionnaire, it conveyed interest in and knowledge about what it meant to be lesbian: "One doctor I went to actually asked on the questionnaire, were you heterosexual, lesbian, gay, or bisexual. And that was wonderful. I could tell that they wouldn't be asking if they weren't open and aware. I thought, this tells them what my lifestyle is, and what my needs are, to a large extent."

Provider characteristics. The second subtheme within this theme was provider characteristics. Participants thought that women providers in general, and lesbians in particular, might have the attitudes and communication skills that were ideal and would be more attuned to the participant's experience: "I like to be sitting across from someone who is more like myself. I feel like they're a lot more in touch with the things that I might be prone to because of my ethnicity as well my sexual preference." However, they cited difficulty with access to lesbian providers: "I'd love to have a lesbian doctor, but on most of the insurance plans that I've had in the past, I've never had access to one."

Not all participants shared these sentiments. A number of women had had good experiences with male providers and nonlesbian providers. Overall, participants thought that the attitude and communication skills of the provider were most important. "To me, the gender or sexual orientation of the caregiver doesn't really matter as long as they treat me the way I think I should be treated, and as long as I feel that they have expertise." Knowledge, or expertise, was another provider characteristic considered ideal. Participants desired providers who were knowledgeable about gender differences in medical concerns, lifestyle choices such as vegetarianism, integrative therapies, and issues that predominantly affect women, such as domestic and sexual abuse: "Healthcare providers [should] have

sort of a broad-based education about the lesbian lifestyle. This might sound basic but yes, lesbians do have children, some lesbians sleep with men occasionally. So, you can't say, they're lesbians so you don't ask them about these issues either."

The lack of evidence-based, health-related information specific to lesbians was cited as a barrier. Women called for more provider education and more research specific to lesbian health: "I know the information is not very clear because they haven't done a lot of research on woman-to-woman transmission [of infectious diseases]." "With some providers, as the lesbian, you're educating them."

Knowing the patient as an individual. The third subtheme was that women wanted a provider who knew them as individuals. Knowledge of the patient's past history and the sense that the provider had read the chart was important. A provider who took a personal interest in the "whole person" and with whom a patient could have more of a "peer relationship" was discussed as ideal more often among the younger women: "I wish when I went to the doctor I was going for more than [pelvic examinations] because I've never really had a doctor that I wanted to talk to about my life, or my lifestyle choices, or anything like that, and I'm really just going because someone told me I needed a Pap smear every year."

Women found that providers assumed they were heterosexual, and this meant that they had to repeatedly disclose their sexual orientation: "But every single time [when asked about birth control needs] you have to explain, and you have to come out, it's like, here we go again." This led to avoidance of healthcare (to avoid repeated disclosure) or nondisclosure and perhaps unnecessary tests and examinations.

DISCUSSION

Our focus group analysis revealed that participants preferred care that is comprehensive in scope, person centered, and inclusive of them as lesbians; often their actual experiences were otherwise. Comprehensive care, as formulated by participants, follows the biopsychosocial model of treating the whole person[27] and includes, or at least accepts as legitimate, integrative medicine practices. Women perceived care as inclusive of them as lesbians if it did not assume heterosexuality and was not discriminatory.

Comprehensive care and person-centered care emerged as interwoven ideas in our focus group analyses. Women favored integration of all aspects of their healthcare (including preventive, nutritional, acute, medical subspecialty, reproductive, gynecological, and mental healthcare). To deliver such comprehensive care, women stressed that providers must first acknowledge them as individuals (not widgets) and take the time to listen attentively to their specific needs. Unfortunately, women noted a gap between actual and ideal care; the actual care they received was often perceived as fragmented and shaming.

To our knowledge, the structure of healthcare delivery preferred by lesbians has not been explored previously. However, prior surveys of lesbians examining the doctor-patient relationship have identified a preference for providers who are compassionate and supportive, consistent with a biopsychosocial approach.[28] This supports our finding that in terms of doctor-patient interactions, the ideal providers have excellent communication skills and approach their patients with a nonjudgmental, accepting attitude about a range of sensitive issues, including obesity, smoking, and sexual orientation. Whereas others have suggested that lesbians tend to prefer female providers,[29] in our study, providers' communication attributes and holistic approach were more important to participants than provider gender. This is consistent with an emerging literature suggesting that, for many women, physician gender is not a dominant issue.[30–32]

The preferences and experiences of healthcare described by the participants were consistent with the documented preferences among women in general in that they experienced healthcare at the level of the model of care, the process of care, and the provider of care. Consistent with the findings reported by Anderson et al.[22] of the National Study of Women, our findings suggest that women prefer healthcare that is integrated, patient centered, easy to access, confidential, and provided in a comfortable setting by a knowledgeable and attentive provider. Participants in our study experienced healthcare not only as women, however, but also as lesbians. Our results offer insight into the specific preferences of lesbians and the experiences that form them. To treat lesbian health simply as women's health would be

to neglect important concerns raised by our participants and echoed in the lesbian health literature.

In general, the participants typically experienced healthcare as exclusive of them as lesbians and related this to heterosexism. Heterosexism is manifest in an implicit assumption of heterosexuality and in discriminatory treatment once lesbian orientation is disclosed. The concepts of assumed heterosexuality and discrimination against homosexuals are well described in the literature.[15,28,33] However, our intensive qualitative approach allowed us to understand some of the determinants of lesbians' perception of heterosexism at the provider level, at the system level, and at the level of scientific inquiry. That is, participants operationalized the construct of heterosexism.

First, at the provider level, participants reported experiencing assumed heterosexuality: providers frame their history taking in heterosexual terms (e.g., asking about birth control use or possible pregnancy). Neutral questions including gay, lesbian, or bisexual orientation as response options were perceived as inclusive. Likewise, a trusting patient-provider relationship promoted disclosure of sexual orientation. Even after disclosing sexual orientation, participants described unintentional discrimination, stemming from providers' lack of knowledge about lesbian health issues (e.g., STD prevention, cancer risk), perhaps reflecting in part the relative paucity of clinically oriented reviews of these topics in the primary care literature.[10] They also experienced or at least feared direct discrimination in the form of scorn ("that crazy lesbian") or remoteness (e.g., shifting abruptly to the next topic). As has been shown in other studies, these experiences led some women to avoid care, potentially jeopardizing their health.[8] In contrast, the opportunity for disclosure and acceptance of sexual orientation led to a perception of better healthcare. Prior studies suggest that a strong patient-provider relationship can indeed lead to better healthcare outcomes.[34–36]

Second, participants experienced the entire healthcare delivery system as being heterosexist. Assumed heterosexuality was evident in waiting rooms lacking lesbian-specific health information and health forms often asking questions about marital status without an option for specifying lesbian orientation. Lesbians feared direct discrimination if their sexual orientation was discovered by staff through chart review. Confidentiality of medical records was considered essential. In the current political climate where confidentiality of health information is receiving unprecedented scrutiny,[37,38] our findings clarify that the issue is of paramount importance to lesbians.

Third, on a global level, women identified the lack of scientific inquiry into questions specific to the heath needs of lesbians as devaluing them and compromising their health. Participants argued that this should be rectified through further research, consistent with the IOM's call for attention to this field.[1]

This study has several limitations. By design, a small sample was recruited from the Boston area, so results may not be generalizable to lesbians living in other geographic regions. Results may also not generalize to heterosexual women (especially the lesbian-specific findings), although some of the priorities of women in our sample mirror those of general populations of women participating in parallel studies.[22,39] Results also may not be generalizable to elderly lesbians; the oldest participant in our study was 63 years old. Another limitation is that participants self-identified as lesbian; their responses may not represent the views of lesbians who would be uncomfortable disclosing their sexual orientation in any setting, even a confidential focus group.

This study also had several strengths. Given the dearth of information about lesbians' preferences for healthcare,[1] the inductive nature of qualitative research[40–42] commends it to this type of question. Unlike many studies of lesbians, our study successfully recruited a sample with diverse age range, race and ethnicity, and access to healthcare. This is consistent with the reality that lesbians come from all demographic and socioeconomic backgrounds.[1]

In conclusion, it is clear that although lesbians do not always disclose their sexual orientation to providers, they do seek healthcare and are likely part of most practices that include women. This has important implications for healthcare providers, for institutions, and for society. Healthcare providers should ask about sexual orientation in a neutral way. They should respond to a patient's disclosure of sexual orientation with an accepting, person-centered attitude and with evidence-based healthcare information that takes all aspects of the patient's health into account. Insti-

tutions seeking to outreach to lesbians should develop systems of care that integrate primary care, mental healthcare, and integrative healthcare. Institutions should also ensure that all staff that interface with patients and their families approach sexual orientation sensitively and confidentially. Indeed, even the physical environment should be welcoming, with lesbian-specific pamphlets in the waiting room and health information forms inclusive of lesbians. At a societal level, more research is needed on lesbian-specific healthcare issues so that providers can be trained to deliver evidence-based care. Based on our participants' input, such measures would be expected to enhance the healthcare experience, and perhaps even healthcare outcomes, of lesbians.

ACKNOWLEDGMENTS

We gratefully acknowledge the technical expertise provided by Bonnie J. Sherman, Ph.D., and Valerie W. Jackson, B.A.

REFERENCES

1. Solarz A, ed. Institute of Medicine. Lesbian health: Current assessment and directions for the future. Washington, DC: National Academy Press, 1999.
2. Gonsiorek J, Weinrich J. Homosexuality: Research implications for public policy. Newbury Park, CA: Sage Publications, 1991.
3. Aaron DJ, Chang YF, Markovic N, LaPorte RE. Estimating the lesbian population: A capture-recapture approach. J Epidemiol Community Health 2003;57: 207.
4. Rankow EJ. Breast and cervical cancer among lesbians. Womens Health Issues 1995;5:123.
5. Diamant AL, Schuster MA, McGuigan K, Lever J. Lesbians' sexual history with men: Implications for taking a sexual history. Arch Intern Med 1999;159:2730.
6. Diamant AL, Schuster MA, Lever J. Receipt of preventive health care services by lesbians. Am J Prev Med 2000;19:141.
7. Diamant AL, Wold C, Spritzer K, Gelberg L. Health behaviors, health status, and access to and use of health care: A population-based study of lesbian, bisexual, and heterosexual women. Arch Fam Med 2000;9:1043.
8. Rankow E, Tessaro I. Cervical cancer risk and Papanicolaou screening in a sample of lesbian and bisexual women. J Fam Pract 1998;47:139.
9. Bradford J, Ryan C. The National Lesbian Health Care Survey: Final report. Washington, DC: National Lesbian and Gay Health Foundation, 1987.
10. White JC, Levinson W. Lesbian health care. What a primary care physician needs to know. West J Med 1995;162:463.
11. Abbott LJ. The use of alcohol by lesbians: A review and research agenda. Subst Use Misuse 1998;33:2647.
12. Johnson SR, Smith EM, Guenther SM. Comparison of gynecologic health care problems between lesbians and bisexual women. A survey of 2,345 women. J Reprod Med 1987;32:805.
13. Smith EM, Johnson SR, Guenther SM. Health care attitudes and experiences during gynecologic care among lesbians and bisexuals. Am J Public Health 1985;75:1085.
14. Carroll NM. Optimal gynecologic and obstetric care for lesbians. Obstet Gynecol 1999;93:611.
15. Trippet SE, Bain J. Physical health problems and concerns of lesbians. Women Health 1993;20:59.
16. Johnson SR, Guenther SM, Laube DW, Keettel WC. Factors influencing lesbian gynecologic care: A preliminary study. Am J Obstet Gynecol 1981;140:20.
17. White JC, Dull VT. Health risk factors and health-seeking behavior in lesbians. J Womens Health 1997;6:103.
18. Asbury J. Overview of focus group research. Qual Health Res 1995;5:414.
19. Morgan DL. Why things (sometimes) go wrong in focus groups. Qual Health Res 1995;5:516.
20. Scholle SH, Weisman CS, Anderson R, Weitz T, Freund KM, Binko J. Women's satisfaction with primary care: A new measurement effort from the PHS National Centers of Excellence in Women's Health. Womens Health Issues 2000;10:1.
21. Poses RM, Isen AM. Qualitative research in medicine and health care: Questions and controversy. J Gen Intern Med 1998;13:32.
22. Anderson RT, Barbara AM, Weisman C, et al. A qualitative analysis of women's satisfaction with primary care from a panel of focus groups in the National Centers of Excellence in Women's Health. J Womens Health Gend Based Med 2001;10:637.
23. Denzin N, Lincoln Y. Handbook of qualitative research, 2nd ed. Thousand Oaks, CA: Sage Publications, 2000.
24. Glaser B, Strauss A. The discovery of grounded theory: The strategies for qualitative research. Hawthorne, NY: Aldine Transaction, 1967.
25. Ryan GW, Bernard HR. Techniques to identify themes. Field Methods 2003;15:85.
26. Waitzkin H. On studying the discourse of medical encounters. A critique of quantitative and qualitative methods and a proposal for reasonable compromise. Med Care 1990;28:473.
27. Engel GL. The need for a new medical model: A challenge for biomedicine. Science 1977;196:129.
28. Stevens PE. Lesbians' health-related experiences of care and noncare. West J Nurs Res 1994;16:639.
29. Zeidenstein L. Gynecological and childbearing needs of lesbians. J Nurse Midwifery 1990;35:10.
30. Plunkett BA, Kohli P, Milad MP. The importance of physician gender in the selection of an obstetrician or a gynecologist. Am J Obstet Gynecol 2002;186:926.

31. Fisher WA, Bryan A, Dervaitis KL, Silcox J, Kohn H. It ain't necessarily so: Most women do not strongly prefer female obstetrician-gynaecologists. J Obstet Gynaecol Can 2002;24:885.
32. Howell EA, Gardiner B, Concato J. Do women prefer female obstetricians? Obstet Gynecol 2002;99:1031.
33. Stevens PE, Hall JM. Stigma, health beliefs and experiences with health care in lesbian women. Image J Nurs Sch 1988;20:69.
34. Safran DG, Taira DA, Rogers WH, Kosinski M, Ware JE, Tarlov AR. Linking primary care performance to outcomes of care. J Fam Pract 1998;47:213.
35. Kaplan SH, Greenfield S, Ware JE Jr. Assessing the effects of physician-patient interactions on the outcomes of chronic disease [published erratum appears in Med Care 1989;27:679]. Med Care 1989;27:S110.
36. McLeod JR. Accepting homosexuality. Can Fam Physician 1994;40:1505.
37. Welch CA. Sacred secrets—The privacy of medical records. N Engl J Med 2001;345:371.
38. Annas GJ. A national bill of patients' rights. N Engl J Med 1998;338:695.
39. Bean-Mayberry BA, Chang CC, McNeil MA, Whittle J, Hayes PM, Scholle SH. Patient satisfaction in women's clinics versus traditional primary care clinics in the Veterans Administration. J Gen Intern Med 2003;18:175.
40. Stevens PE. Focus groups: Collecting aggregate-level data to understand community health phenomena. Public Health Nurs 1996;13:170.
41. Mishler EG. Research interviewing: Context and narrative. Cambridge, MA: Harvard University Press, 1991.
42. Berkwits M, Inui TS. Making use of qualitative research techniques. J Gen Intern Med 1998;13:195.

Address reprint requests to:
Margaret R. Seaver, M.D., M.P.H.
VA Boston Healthcare System
150 S. Huntington Avenue (11PCC/WHC)
Boston, MA 02130

E-mail: margrsea@massmed.org

ARTICLE

Local Restaurant Smoking Regulations and the Adolescent Smoking Initiation Process

Results of a Multilevel Contextual Analysis Among Massachusetts Youth

Michael Siegel, MD, MPH; Alison B. Albers, PhD; Debbie M. Cheng, ScD; William L. Hamilton, PhD; Lois Biener, PhD

Objective: To assess whether smoke-free restaurant laws influence the progression from (1) never smoking to early experimentation and (2) early experimentation to established smoking.

Design: A longitudinal, 4-year, 3-wave study of a representative sample of Massachusetts youth.

Setting: A total of 301 Massachusetts communities.

Participants: Study participants were 3834 Massachusetts youths aged 12 to 17 years at baseline, from January 2, 2001, to June 18, 2002, of whom 2791 (72.8%) were reinterviewed after 2 years (from January 30, 2003, to July 31, 2004) and 2217 (57.8%) were reinterviewed after 4 years (from February 16, 2005, to March 26, 2006). Wave 3 respondents were recruited from both those who responded at wave 2 and those who did not.

Main Exposure: The primary predictor of interest is the strength of the local restaurant smoking regulation in the respondents' town of residence at the baseline of each transition period.

Main Outcome Measures: (1) Overall progression to established smoking (having smoked ≥100 cigarettes in one's lifetime), (2) transition from nonsmoking (never having puffed a cigarette) to experimentation, and (3) transition from experimentation to established smoking.

Results: Youths living in towns with a strong restaurant smoking regulation at baseline had significantly lower odds of progressing to established smoking (odds ratio, 0.60; 95% confidence interval, 0.42-0.85) compared with those living in towns with weak regulations. The observed association between strong restaurant smoking regulations and impeded progression to established smoking was entirely due to an effect on the transition from experimentation to established smoking (odds ratio, 0.53; 95% confidence interval, 0.33-0.86).

Conclusion: Local smoke-free restaurant laws may significantly lower youth smoking initiation by impeding the progression from cigarette experimentation to established smoking.

Arch Pediatr Adolesc Med. 2008;162(5):477-483

Author Affiliations:
Departments of Social and Behavioral Sciences (Drs Siegel and Albers) and Biostatistics (Dr Cheng), Boston University School of Public Health, Boston; CNU Associates, Lincoln (Dr Hamilton); and Center for Survey Research, University of Massachusetts, Boston (Dr Biener), Massachusetts.

RECENT EVIDENCE SUGGESTS that laws that protect nonsmokers from secondhand smoke[1-3] by eliminating smoking in restaurants[3-6] may not only protect restaurant workers and customers from secondhand smoke[7,8] but also reduce adolescent smoking initiation[9] by changing the perceived prevalence and social acceptability of smoking among youth.[10] Existing evidence that smoke-free laws reduce youth smoking derives from cross-sectional studies.[11-16] However, the results of the first longitudinal investigation of the impact of local smoke-free restaurant laws on smoking initiation were recently reported.[9] During a 2-year follow-up, Massachusetts youth who lived in a town with a complete restaurant smoking ban had less than half the odds of progressing to established smoking. We report the final results of this study, which reflect the continued follow-up of more than 2000 youths for 4 years. In addition to the extended follow-up, which improves the validity and power of our analysis, this article adds to the literature by addressing a new research question: if smoke-free restaurant laws reduce smoking initiation, do they do so by inhibiting experimentation with cigarettes or by impeding the progression from experimentation to regular smoking?

Although an abundance of literature has examined risk factors for smoking initiation,[17-19] few studies[20] have differentiated factors that influence experimentation from those that influence the progression from experimentation to regular smoking. Yet understanding this difference is critical. It would allow us to determine the age and stage at which youths

are most sensitive to various types of interventions, thus enabling the more specific tailoring and more effective delivery of smoking prevention interventions.

The present study overcomes several important limitations of the existing research. First, most previous research has not included community-level influences on smoking initiation.[21] We are not aware of any previous studies that have examined the impact of smoke-free laws on smoking stage transitions among youths. Second, although most previous community intervention studies have been based on a few communities, we have access to data representing individuals from more than 300 different towns. Third, much of the previous literature is based on cross-sectional designs or on longitudinal designs with only 2 successive observations for each individual (and usually for only a 1- to 2-year period). Our study is a longitudinal analysis that follows up a cohort of youths during a 4-year period, with 3 successive observations for each individual. In summary, our study design provides a unique opportunity to examine not only predictors of overall progression to established smoking but also the specific predictors of transitioning from never smoking to experimentation and from experimentation to established smoking using an integrated multilevel model of adolescent smoking trajectories that incorporates both individual and contextual (community-level) forces.[20-22]

METHODS

SAMPLE

Between January 2, 2001, and June 18, 2002, the Center for Survey Research, University of Massachusetts, obtained a probability sample of 3834 Massachusetts youths, aged 12 to 17 years, by random-digit dialing.[7-10] Of households in which screening interviews were completed, parental permission was obtained to interview 75.9% of eligible youths, and interviews were completed with 84.7% of those.

Between January 30, 2003, and July 31, 2004, we attempted to reinterview all 3834 of the youths in the baseline sample. Interviews were completed with 2791 individuals, for a follow-up rate of 72.8%. Between February 16, 2005, and March 26, 2006, we attempted to reinterview all 2791 youths who were successfully followed up to wave 2 and all youths lost to follow-up at wave 2. Of the former group, 2045 youths were successfully reinterviewed, and of the latter group, 172 youths were successfully reinterviewed, for a total wave 3 sample size of 2217 (57.8% of the baseline sample). The research protocol was approved by the institutional review boards of the University of Massachusetts and Boston University Medical Center.

MAIN OUTCOME MEASURES

Town of Residence

Town of residence at each wave was obtained using the reported zip code. Most (95.6%) of the reinterviewed youths at wave 2 lived in the same town at baseline and 2-year follow-up; 2.7% moved within Massachusetts, and 1.7% moved out of state. Of youths interviewed at wave 3, 91.4% lived in the same town at baseline, 5.6% had moved within Massachusetts, and 3.0% had moved out of state.

Strength of Local Restaurant Smoking Regulation

The strength of the local restaurant regulation in effect in each respondent's town of residence on the date of his or her baseline interview was categorized[23] as follows: (1) strong regulations, no smoking allowed in restaurants and no variances allowed; (2) medium regulations, smoking restricted to enclosed separately ventilated areas or no smoking allowed but variances allowed; and (3) weak regulations, smoking restricted to designated areas or not restricted.

Stages of Smoking Initiation

According to the work of Pierce et al,[24] we defined progression to established smoking as having smoked 100 or more cigarettes in one's lifetime. This measure has been formally validated[25-27] and used in previous studies.[24-31] The experimentation stage of smoking was then defined as the period from trying a cigarette until becoming an established smoker. Thus, the 3 stages of smoking initiation were (1) nonsmoking, (2) experimentation (having tried a cigarette but not smoked 100 cigarettes), and (3) established smoking (having smoked ≥100 cigarettes).

In the first set of analyses, we model overall progression to established smoking (from either nonsmoking or experimentation). In a second set of analyses, we model each of the 2 possible transitions in smoking stages: (1) from nonsmoking to experimentation and (2) from experimentation to established smoking.

Individual-Level Predictor Variables

We examined the effect of the following individual-level baseline variables: (1) age group (12-14, 15-17, and 18-21 years), (2) sex, (3) race (non-Hispanic white vs other), (4) presence of at least 1 adult smoker in the household, (5) presence of at least 1 close friend who smokes, (6) education level of household informant (college graduate or not), (7) annual household income (≤$50 000 vs >$50 000), (8) exposure to antismoking messages at school (yes or no), and (9) self-reported baseline smoking status (nonsusceptible nonsmoker, susceptible nonsmoker, puffer, experimenter, or current smoker). Although none of the study participants included in the analysis had smoked 100 cigarettes at baseline, some had experimented with cigarettes. We controlled for individuals' baseline predisposition to smoking by including in the analysis a set of indicator variables that reflect 5 categories of baseline smoking status: nonsusceptible nonsmoker, susceptible nonsmoker, puffer, experimenter, and current smoker. For this purpose, *nonsmokers* were defined as respondents who had never puffed on a cigarette, *puffers* as those who had puffed but not smoked a whole cigarette, *experimenters* as those who had smoked at least 1 whole cigarette but none within the past 30 days, and *current smokers* as those who had smoked at least 1 cigarette, including 1 or more within the past 30 days. Nonsmokers were further classified based on a measure of susceptibility to smoking that has been shown to reliably predict progression to established smoking.[24-27,29] Nonsmokers were classified as nonsusceptible to smoking if they answered no to the question, "Do you think that you will try a cigarette soon?" and definitely not to the questions, "If one of your best friends were to offer you a cigarette, would you smoke it?" and "At any time during the next year do you think you will smoke a cigarette?"

Town-Level Predictor Variables

We examined the effect of the following town-level variables (included as continuous variables except where noted): (1) the percentage of each town's voters who voted yes on question 1, a 1992

ballot initiative that increased the cigarette tax and created a statewide tobacco control program; (2) the percentage of white residents in each town; (3) the percentage of youths (aged < 18 years) in each town; and (4) town population (< 20 000, 20 000-50 000, and > 50 000). Of many town-level factors examined, these were most strongly related to the strength of local restaurant smoking regulations in Massachusetts towns.[32] The percentage of yes votes on question 1 served as a measure of the baseline level of antismoking sentiment in each town before the proliferation of local restaurant smoking regulations, which correlates with the level of education in the town.[33] All town-level variables were obtained from the 2000 US Census, except for the question 1 vote, which was obtained from the Division of Elections within the Massachusetts office of the secretary of state.

DATA ANALYSIS

Our data set has clustering at 2 levels. First, observations are clustered within individual respondents. Each respondent may contribute up to 2 observations in the data set. Second, respondents are clustered within towns. Because observations among individuals and respondents from the same town may be more similar than observations across respondents or respondents from different towns, we used a multilevel (hierarchical) logistic regression model to examine the relationship between strength of town restaurant smoking regulations at baseline and smoking progression. This procedure accounts for correlation of data within individuals and within town "clusters," reducing the probability of a type I error that could be introduced if this correlation were ignored.[22,34]

All town-level variables were time-independent and assessed at the start of the study (modeled at level 3), except strength of restaurant smoking regulation, which was modeled as a time-varying variable (at level 1), updated at each time point (based on the strength of local restaurant smoking regulation on the interview date). Time-independent individual-level variables (entered at level 2) were sex, race, informant education level, and household income. The following individual-level variables were modeled as time varying (at level 1): age group, presence of a household smoker, presence of a close friend who smokes, exposure to school-based antismoking messages, and baseline smoking status.

Three separate models were fit. The first model assessed overall progression to established smoking (from either nonsmoking or experimentation). The second model assessed progression to experimentation from a nonsmoking stage. The third model assessed progression to established smoking only from experimentation. Additional exploratory analyses stratified by baseline age were conducted to assess whether age moderates the influence of smoke-free restaurant laws on smoking initiation. All analyses were conducted using 2-sided tests and a significance level of .05. Analyses were conducted using HLM statistical software, version 6.0 (Scientific Software International Inc, Lincolnwood, Illinois).

For the baseline sample, survey weights were computed that adjusted for the number of telephones per household and, hence, for the probability of selection, and for nonresponse. The most important differences between respondents who were followed up successfully and those who were lost to attrition were as follows: respondents who followed up were more likely to be younger, to not have a smoker in the household, to have higher household income, to have more highly educated parents, to be never smokers, to not be black or Hispanic, and to not have a close friend who smokes. We created adjustments to the baseline weights by using an iterative ranking procedure[35] that yielded distributions on age, race, smoking status, parental smoking, and parental educational level that either were identical to those at baseline or differed by at most 1 percentage point.

RESULTS

Our study sample consisted of 2791 unique individuals, contributing 4596 observations. Analysis of overall progression to established smoking was based on all 4596 observations (wave 1 to wave 2: 2623; wave 2 to wave 3: 1818; and wave 1 to wave 3: 155). Analyses of separate smoking stage transitions were based on 4491 observations (wave 1 to wave 2: 2572; wave 2 to wave 3: 1768; and wave 1 to wave 3: 151). Sample sizes for the analyses of separate smoking stage transitions were slightly smaller than for overall progression to established smoking because of missing or inconsistent data on smoking stage for some individuals.

BASELINE CHARACTERISTICS OF THE SAMPLE

Of our total sample of 4596 observations (transitions or nontransitions from nonsmoking or experimentation to established smoking), the overall rate of progression to established smoking during the follow-up periods was 9.3%, and it varied from 9.6% and 9.8% for youths living in towns with weak and medium regulations, respectively, to 7.9% for youths living in towns with strong local restaurant smoking regulations (**Table 1**).

PREDICTORS OF OVERALL PROGRESSION TO ESTABLISHED SMOKING

No association was found between medium restaurant smoking regulations and progression to established smoking (odds ratio [OR], 0.93; 95% confidence interval [CI], 0.67-1.30) (**Table 2**). However, youths living in towns with a strong restaurant smoking regulation at baseline had significantly lower odds of progressing to established smoking (OR, 0.60; 95% CI, 0.42-0.85) compared with those living in towns with weak regulations.

Other significant predictors of increased odds of progression to established smoking included older age group (OR, 2.02; 95% CI, 1.16-3.51 [for youths aged 18-21 years at baseline]), previous experimentation with cigarettes, presence of an adult smoker in the household (OR, 1.54; 95% CI, 1.20-1.99), presence of a close friend who smokes (OR, 1.91; 95% CI, 1.45-2.53), being male (OR, 0.66; 95% CI, 0.51-0.85 [for females]), and living in a town with more white residents (Table 2).

The association between strong restaurant smoking regulations and overall progression to established smoking seemed to differ by age of the respondent at baseline. The association was present for young (aged 12-14 years) individuals (OR, 0.63) and middle-aged (aged 15-17 years) individuals (OR, 0.52), but not for older (aged 18-21 years) individuals (OR, 1.17).

PREDICTORS OF TRANSITION FROM NONSMOKING TO EXPERIMENTATION

The strength of the local restaurant smoking regulation was not significantly associated with the transition from nonsmoking to experimentation (OR, 1.18; 95% CI, 0.94-1.49 [for strong regulations]) (Table 2). Important pre-

Table 1. Baseline Characteristics of Cohort and Progression to Established Smoking by Individual and Contextual Variables

Variable	Total, No. (%)[a,b]	Progression to Established Smoking, %[a,c]
Total	4596	9.3
Main predictor variable (level 1)		
Strength of local restaurant smoking regulation		
Weak	2529 (55.2)	9.6
Medium	1049 (22.8)	9.8
Strong	1018 (22.0)	7.9
Individual-level time-varying covariates (level 1)		
Age group, y		
12-14	1832 (40.5)	5.7
15-17	2303 (49.6)	11.2
18-21	461 (9.9)	14.1
Baseline smoking status		
Nonsusceptible never smoker	2664 (58.3)	2.7
Susceptible never smoker	769 (16.8)	7.1
Puffed	557 (12.0)	13.5
Smoked whole cigarette	406 (8.8)	30.0
Smoked in past 30 d	200 (4.1)	53.9
Presence of adult smoker in household		
No	3243 (70.4)	7.4
Yes	1353 (29.6)	13.7
Presence of close friend who smokes		
No	3240 (71.0)	5.1
Yes	1356 (29.0)	19.5
Exposure to school-based antismoking messages		
No	1331 (32.2)	9.7
Yes	2795 (67.8)	8.6
4-y Follow-up period (wave 1 to wave 3)		
No	4441 (96.6)	8.9
Yes	155 (3.4)	20.1
Individual-level covariates (level 2)		
Sex		
Male	2318 (50.9)	10.1
Female	2278 (49.1)	8.4
Race/ethnicity		
Non-Hispanic white	3788 (82.3)	9.3
Other	775 (17.7)	8.4
Informant education level		
Not college graduate	2342 (52.9)	10.3
College graduate	2173 (47.1)	8.2
Annual household income, $		
≤50 000	945 (26.4)	9.6
>50 000	2813 (73.6)	8.9
Town-level covariates (level 3)		
Yes vote on question 1, %[d]		
≤50	2390 (53.1)	9.2
>50	2206 (46.9)	9.3
Residents who are white, %[d]		
≤90	1645 (35.8)	7.7
>90	2951 (64.2)	10.1
Residents who are youths, %[d]		
≤25	2317 (50.8)	9.5
>25	2279 (49.2)	9.0
Town population		
<20 000	1833 (39.2)	10.6
20 000-50 000	1656 (36.6)	8.8
>50 000	1107 (24.3)	7.7

[a] Percentages in table are weighted to reflect initial probability of participant selection into sample.
[b] Total number of observations (individuals may have ≤2 follow-up observations).
[c] Progression to established smoking is defined as having smoked 100 cigarettes in one's lifetime.
[d] Modeled as continuous variables in analysis.

dictors of transition to experimentation included age (OR, 2.04 [for youths aged 15-17 years] and 2.67 [for youths aged 18-21 years]), susceptibility to smoking (OR, 2.78 for susceptible youths), presence of an adult smoker in the household (OR, 1.47), presence of a close friend who smokes (OR, 1.63), and parental education level (OR, 0.80 for college graduates) (Table 2). No effect was found of sex or the proportion of white residents in the respondent's town.

PREDICTORS OF TRANSITION FROM EXPERIMENTATION TO ESTABLISHED SMOKING

Although living in a town with a medium-strength restaurant smoking regulation had no significant association with the transition from experimentation to established smoking, youths living in towns with strong regulations had significantly lower odds of making this transition (OR, 0.53; 95% CI, 0.33-0.86) (Table 2). Other important predictors of transition from experimentation to established smoking included age (OR, 2.47 for youths aged 18-21 years at baseline), more advanced experimentation, presence of a close friend who smokes (OR, 1.85), sex (OR, 0.61 for females), and percentage of white residents in one's town (OR, 1.38 for each 10–percentage point increase). The presence of an adult smoker in the household had no significant effect on this transition. Informant education level was also not a significant predictor of the transition from experimentation to established smoking.

The association between strong restaurant smoking regulations and progression from experimentation to established smoking seemed to differ by age of the respondent at baseline. The association was present for young (aged 12-14 years) individuals (OR, 0.57) and middle-aged (aged 15-17 years) individuals (OR, 0.60) but not for older (aged 18-21 years) individuals (OR, 0.99).

COMMENT

To our knowledge, this is the first study to examine the effect of restaurant smoking laws on different stages of the smoking initiation process. Using a hierarchical repeated-measures model that examined individual and contextual factors that influence smoking initiation, we found not only that strong restaurant smoking regulations were associated with a significant decrease in the odds of progression to established smoking among youths but also that this association was specific to the transition from experimentation to established smoking. In addition, the effects of the smoking regulations seemed to be stronger for young and middle-aged youths (aged 12-17 years).

The analysis did not directly examine the mechanisms by which smoking bans might reduce smoking initiation. Nevertheless, the findings are consistent with the conceptual hypothesis, from a previous study,[7] that restaurant smoking bans affect smoking initiation by (1) reducing youths' exposure to smokers in public places, which lowers their perception of smoking prevalence, and (2) changing the perceived social acceptability of smoking. Both of these effects would be expected to influence

Table 2. Data for Overall Progression to Established Smoking, Transition From Nonsmoking to Experimentation, and Transition From Experimentation to Established Smoking

Variable	Overall Progression to Established Smoking[a,b]	Transition From Nonsmoking to Experimentation[a,c]	Transition From Experimentation to Established Smoking[a,d]
Main predictor variable (level 1)			
Strength of local restaurant smoking regulation			
Weak	1 [Reference]	1 [Reference]	1 [Reference]
Medium	0.93 (0.67-1.30)	1.01 (0.78-1.31)	0.78 (0.47-1.30)
Strong	0.60 (0.42-0.85)	1.18 (0.94-1.49)	0.53 (0.33-0.86)
Individual-level time-varying covariates (level 1)			
Age group, y			
12-14	1 [Reference]	1 [Reference]	1 [Reference]
15-17	1.22 (0.90-1.64)	2.04 (1.63-2.54)	1.09 (0.65-1.82)
18-21	2.02 (1.16-3.51)	2.67 (1.51-4.72)	2.47 (1.04-5.86)
Baseline smoking status			
Nonsusceptible never smoker	1 [Reference]	1 [Reference]	1 [Reference]
Susceptible never smoker	2.84 (1.85-4.36)	2.78 (2.11-3.66)	NA
Puffed	4.65 (3.16-6.84)	NA	NA
Smoked whole cigarette	14.60 (9.84-21.7)	NA	3.14 (2.07-4.76)
Smoked in past 30 d	45.80 (29.10-72.20)	NA	11.50 (7.36-18.10)
Presence of adult smoker in household			
No	1 [Reference]	1 [Reference]	1 [Reference]
Yes	1.54 (1.20-1.99)	1.47 (1.16-1.87)	1.31 (0.87-1.99)
Presence of close friend who smokes			
No	1 [Reference]	1 [Reference]	1 [Reference]
Yes	1.91 (1.45-2.53)	1.63 (1.25-2.13)	1.85 (1.24-2.77)
Exposure to school-based antismoking messages			
No	1 [Reference]	1 [Reference]	1 [Reference]
Yes	0.76 (0.56-1.04)	0.94 (0.75-1.18)	0.74 (0.47-1.16)
4-y Follow-up period (wave 1 to wave 3)			
No	1 [Reference]	1 [Reference]	1 [Reference]
Yes	2.92 (1.58-5.41)	1.82 (1.14-2.92)	3.88 (1.42-10.60)
Individual-level covariates (level 2)			
Sex			
Male	1 [Reference]	1 [Reference]	1 [Reference]
Female	0.66 (0.51-0.85)	0.97 (0.78-1.21)	0.61 (0.42-0.90)
Race/ethnicity			
Non-Hispanic white	1 [Reference]	1 [Reference]	1 [Reference]
Other	1.09 (0.66-1.78)	0.91 (0.67-1.23)	1.33 (0.77-2.30)
Informant education level			
Not college graduate	1 [Reference]	1 [Reference]	1 [Reference]
College graduate	0.90 (0.68-1.18)	0.80 (0.65-0.99)	0.79 (0.53-1.18)
Annual household income, $			
≤50 000	1 [Reference]	1 [Reference]	1 [Reference]
>50 000	1.03 (0.74-1.43)	1.08 (0.80-1.47)	1.07 (0.60-1.90)
Town-level covariates (level 3)			
Percentage yes vote on question 1[e]	0.93 (0.80-1.08)	1.13 (0.99-1.29)	0.87 (0.69-1.11)
Percentage of residents who are white[e]	1.24 (1.05-1.46)	1.10 (0.96-1.27)	1.38 (1.13-1.69)
Percentage of residents who are youths[e]	0.91 (0.62-1.33)	0.76 (0.58-0.99)	0.89 (0.51-1.58)
Town population			
<20 000	1 [Reference]	1 [Reference]	1 [Reference]
20 000-50 000	0.86 (0.61-1.21)	0.83 (0.64-1.08)	0.83 (0.49-1.40)
>50 000	0.91 (0.56-1.49)	1.03 (0.68-1.56)	0.81 (0.39-1.71)

Abbreviation: NA, data not applicable.
[a] Data are given as odds ratio (95% confidence interval).
[b] Progression to established smoking is defined as having smoked 100 cigarettes in one's lifetime. Analyses based on 2791 individuals living in 301 towns, contributing 4596 observations.
[c] Analyses based on 2091 individuals living in 286 towns, contributing 3301 observations.
[d] Analyses based on 808 individuals living in 240 towns, contributing 1059 observations.
[e] The odds ratios were associated with each 10–percentage point increase in variable.

the transition from experimentation to established smoking but not experimentation in the first place.

In general, our results are consistent with the finding from earlier studies[17,20,28,36-46] that purely individual-level factors are more important in influencing smoking experimentation, whereas community-level factors mainly exert an influence on the transition from experimentation to regular cigarette use. The major community-

level factor we examined—smoke-free restaurant policies—exerted an effect only on the transition to regular use. Parental smoking and parental education level, however, were significant predictors only for cigarette experimentation.

These results have a number of important public health policy implications. First, they suggest that local smoke-free restaurant laws may decrease youth smoking initiation. These results extend the follow-up period of our preliminary study[9] to 4 years and improve the overall power of our analysis. If it represents a true effect, the observed 40% reduction in the odds of progression to established smoking in towns with local restaurant smoking bans would suggest that smoke-free policies may be the most effective intervention available to reduce youth smoking.

Second, these findings demonstrate the importance of considering individual-level and contextual factors and of separating the effects of each on various stages of the smoking initiation process. We found, for example, that parental smoking is a strong factor in predicting which youths will experiment with cigarettes. However, once that experimentation has taken place, parental smoking was no longer a factor in the progression to regular cigarette use. This suggests that although interventions that address the role of parents in putting youths at risk for smoking may be effective in deterring experimentation, they are unlikely to succeed in halting the progression to regular smoking among youths who have already tried cigarettes.

Third, our results suggest that the early and middle periods of adolescence are critical times of susceptibility to public policy interventions. Smoke-free restaurant regulations seemed to be less effective for youths older than 18 years.

The primary potential threat to the validity of our findings is the relatively high rate of loss to follow-up in the study. Although not unusual for a telephone survey following up individuals for 4 years, the follow-up rates of 72.8% at wave 2 and 57.8% at wave 3 introduce the possibility of a differential loss to follow-up bias. However, we believe loss to follow-up is an unlikely explanation of our results because loss to follow-up was lower among youths living in towns with strong regulations than among youths living in towns with weak regulations, and those lost to follow-up are more likely to be smokers. At wave 2, for example, only 23.9% of youths living in towns with strong regulations at wave 1 were lost to follow-up compared with 29.0% of youths living in towns with weak regulations. In addition, we expect that those lost to follow-up are more likely to be smokers because follow-up was higher for households with a higher education level and income and for youths who were never smokers without household smokers or close friends who smoked. This combination would bias the results toward the null and result in an underestimate of the effect of smoking regulations because it would produce differentially fewer smokers at follow-up among respondents living in towns with weak regulations. Also, sampling weights were used to yield a study sample that was identical to the full baseline sample in terms of age, race, smoking status, and parental smoking and education level.

A second limitation of this research is that it is not clear whether the results are generalizable to other populations. Local clean indoor air regulations in Massachusetts were adopted under a broad and aggressive statewide antismoking campaign that included a state-of-the-art media campaign that aimed to denormalize smoking and educate the public about the hazards of secondhand smoke. It is not clear whether restaurant smoking laws would have similar effects on youth smoking behavior in states where widespread antismoking programs are not present.

Despite these limitations, the evidence presented in this article suggests that local smoke-free restaurant laws, if they are strong enough (ie, complete smoking bans), may significantly lower youth smoking initiation by impeding the progression from cigarette experimentation to regular or established smoking. Further research is needed to examine the effect of local restaurant smoking regulations in other states, especially in the absence of statewide antismoking programs, and to explore potential mediating factors between restaurant smoking bans and youth smoking behavior.

Accepted for Publication: October 23, 2007.
Correspondence: Michael Siegel, MD, MPH, Departments of Social and Behavioral Sciences, Boston University School of Public Health, 715 Albany St, TW2, Boston, MA 02118 (mbsiegel@bu.edu).
Author Contributions: Dr Siegel 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*: Siegel and Biener. *Acquisition of data*: Hamilton. *Analysis and interpretation of data*: Siegel, Albers, and Cheng. *Drafting of the manuscript*: Siegel. *Critical revision of the manuscript for important intellectual content*: Siegel, Albers, Cheng, Hamilton, and Biener. *Statistical analysis*: Siegel, Albers, Cheng, and Hamilton. *Obtained funding*: Siegel and Biener. *Administrative, technical, or material support*: Biener. *Study supervision*: Siegel.
Financial Disclosure: None reported.
Funding/Support: This study was supported by grants from the National Cancer Institute; grant 2R01 CA86257-05 from the State and Community Tobacco Control Interventions Research Grant Program; and the Flight Attendant Medical Research Institute.
Role of the Sponsor: The funding organizations had no role in the design and conduct of the study, collection, management, analysis, and interpretation of data or the preparation, review, and approval of the manuscript.

REFERENCES

1. US Department of Health and Human Services. *The Health Consequences of Involuntary Exposure to Tobacco Smoke: A Report of the Surgeon General.* Atlanta, GA: Coordinating Center for Health Promotion, Centers for Disease Control and Prevention and Office on Smoking and Health, National Center for Chronic Disease Prevention and Health Promotion; 2006.
2. California Environmental Protection Agency. *Proposed Identification of Environmental Tobacco Smoke as a Toxic Air Contaminant.* Sacramento, CA: Air Resources Board, Office of Environmental Health Hazard Assessment; 2005.
3. Siegel M, Barbeau EM, Osinubi OY. The impact of tobacco use and secondhand smoke on hospitality workers. *Clin Occup Environ Med.* 2006;5(1):31-42.
4. American Nonsmokers' Rights Foundation. Municipalities with local 100% smoke-free laws: cumulative number effective by year, 1990-2007. http://www.no-smoke.org/pdf/100ordgraph.pdf. Accessed June 1, 2007.

5. American Nonsmokers' Rights Foundation. Overview list: how many smokefree laws? http://www.no-smoke.org/pdf/mediaordlist.pdf. Accessed June 1, 2007.
6. American Nonsmokers' Rights Foundation. Summary of 100% smokefree state laws and population protected by state and local laws. http://www.no-smoke.org/pdf/SummaryUSPopList.pdf. Accessed June 1, 2007.
7. Siegel M, Albers AB, Cheng DM, Biener L, Rigotti NA. Effect of local restaurant smoking regulations on environmental tobacco smoke exposure among youths. *Am J Public Health*. 2004;94(2):321-325.
8. Albers AB, Siegel M, Cheng DM, Rigotti NA, Biener L. Effects of restaurant and bar smoking regulations on exposure to environmental tobacco smoke among Massachusetts adults. *Am J Public Health*. 2004;94(11):1959-1964.
9. Siegel M, Albers AB, Cheng DM, Biener L, Rigotti NA. Effect of local restaurant smoking regulations on progression to established smoking among youths. *Tob Control*. 2005;14(5):300-306.
10. Albers AB, Siegel M, Cheng DM, Biener L, Rigotti NA. Relation between local restaurant smoking regulations and attitudes towards the prevalence and social acceptability of smoking: a study of youths and adults who eat out predominantly at restaurants in their town. *Tob Control*. 2004;13(4):347-355.
11. Chaloupka FJ, Grossman M. *Price, Tobacco Control Policies and Youth Smoking*. Cambridge, MA: National Bureau of Economic Research; 1996. Working paper 5740.
12. Chaloupka FJ, Pacula RL. Sex and race differences in young people's responsiveness to price and tobacco control policies. *Tob Control*. 1999;8(4):373-377.
13. Gruber J, Zinman J. *Youth Smoking in the US: Evidence and Implications* . Cambridge, MA: National Bureau of Economic Research; 2000. Working paper 7780.
14. Wakefield MA, Chaloupka FJ, Kaufman NJ, Orleans CT, Barker DC, Ruel EE. Effect of restrictions on smoking at home, at school, and in public places on teenage smoking: cross sectional study [published correction appears in *BMJ*. 2000;31(7261):623]. *BMJ*. 2000;321(7257):333-337.
15. Wasserman J, Manning WG, Newhouse JP, et al. The effects of excise taxes and regulations on cigarette smoking. *J Health Econ*. 1991;10(1):43-64.
16. Giovino GA, Tworek C, Chaloupka FJ, et al. Smoke-free air policies and adolescent smoking: United States, 1991-2001. Paper presented at: 12th World Conference on Tobacco or Health; August 5, 2003; Helsinki, Finland.
17. US Department of Health and Human Services. *Preventing Tobacco Use Among Young People: A Report of the Surgeon General*. Atlanta, GA: Office on Smoking and Health, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention, US Dept of Health and Human Services; 1994.
18. Conrad KM, Flay BR, Hill D. Why children start smoking cigarettes: predictors of onset. *Br J Addict*. 1992;87(12):1711-1724.
19. Tyas SL, Pederson LL. Psychosocial factors related to adolescent smoking: a critical review of the literature. *Tob Control*. 1998;7(4):409-420.
20. Mayhew KP, Flay BR, Mott JA. Review: stages in the development of adolescent smoking. *Drug Alcohol Depend*. 2000;59(suppl 1):S61-S81.
21. Wilcox P. An ecological approach to understanding youth smoking trajectories: problems and prospects. *Addiction*. 2003;98(suppl 1):57-77.
22. Raudenbusch SW, Bryk AS. *Hierarchical Linear Models: Applications and Data Analysis Methods*. 2nd ed. Thousand Oaks, CA: Sage Publications; 2002.
23. Skeer M, Siegel M. The descriptive epidemiology of local restaurant smoking regulations in Massachusetts: an analysis of the protection of restaurant customers and workers. *Tob Control*. 2003;12(2):221-226.
24. Pierce JP, Choi WS, Gilpin EA, et al. Tobacco industry promotion of cigarettes and adolescent smoking. *JAMA*. 1998;279(7):511-515.
25. Pierce JP, Choi WS, Gilpin EA, Farkas AJ, Merritt RK. Validation of susceptibility as a predictor of which adolescents take up smoking in the United States. *Health Psychol*. 1996;15(5):355-361.
26. Choi WS, Pierce JP, Gilpin EA, Farkas AJ, Berry CC. Which adolescent experimenters progress to established smoking in the United States. *Am J Prev Med*. 1997;13(5):385-391.
27. Pierce JP, Farkas AJ, Evans N, et al. An improved surveillance measure for adolescent smoking? *Tob Control*. 1995;4(suppl 1):S47-S56.
28. Mowery PD, Farrelly MC, Haviland L, Gable JM, Wells HE. Progression to established smoking among US youths. *Am J Public Health*. 2004;94(2):331-337.
29. Siegel M, Biener L. The impact of an antismoking media campaign on progression to established smoking: results of a longitudinal youth study. *Am J Public Health*. 2000;90(3):380-386.
30. Siegel M, Biener L, Rigotti NA. The effect of local tobacco sales laws on adolescent smoking initiation. *Prev Med*. 1999;29(5):334-342.
31. Choi WS, Ahluwalia JS, Harris KJ, et al. Progression to established smoking: the influence of tobacco marketing. *Am J Prev Med*. 2002;22(4):228-233.
32. Skeer M, George S, Hamilton WL, et al. Town-level characteristics and smoking policy adoption in Massachusetts: are local restaurant smoking regulations fostering disparities in health protection? *Am J Public Health*. 2004;94(2):286-292.
33. Hamilton WL, Biener L, Rodger CN. Who supports tobacco excise taxes? factors associated with towns' and individuals' support in Massachusetts. *J Public Health Manag Pract*. 2005;11(4):333-340.
34. Hedeker D, Gibbons RD, Flay BR. Random-effects regression models for clustered data with an example from smoking prevention research. *J Consult Clin Psychol*. 1994;62(4):757-765.
35. Bishop YM, Fienberg SE, Holland PW. *Discrete Multivariate Analysis: Theory and Practice*. Cambridge, MA: MIT Press; 1975.
36. Stern RA, Prochaska JO, Velicer WF, Elder JP. Stages of adolescent cigarette smoking acquisition: measurement and sample profiles. *Addict Behav*. 1987;12(4):319-329.
37. Prochaska JO, DiClemente CC. Stages and processes of self-change of smoking: toward an integrative model of change. *J Consult Clin Psychol*. 1983;51(3):390-395.
38. Wellman RJ, DiFranza JR, Savageau JA, Dussault GF. Short term patterns of early smoking acquisition. *Tob Control*. 2004;13(3):251-257.
39. Elder JP, DeMoor C, Young RL, Wildy M, Molgaard C, Golbeck A. Stages of adolescent tobacco-use acquisition. *Addict Behav*. 1990;15(5):449-454.
40. Leventhal H, Fleming R, Glynn K. A cognitive-developmental approach to smoking intervention. In: Maes S, Spielberger CD, Defares PB, Sarason IG, eds. *Topics in Health Psychology*. New York, NY: John Wiley & Sons Inc; 1988:79-105.
41. Leventhal H, Cleary PD. The smoking problem: a review of the research and theory in behavioral risk modification. *Psychol Bull*. 1980;88(2):370-405.
42. Shiffman S. Tobacco "chippers": individual differences in tobacco dependence. *Psychopharmacology (Berl)*. 1989;97(4):539-547.
43. McNeill AD. The development of dependence on smoking in children. *Br J Addict*. 1991;86(5):589-592.
44. Hirschman RS, Leventhal H, Glynn K. The development of smoking behavior: conceptualization and supportive cross-sectional survey data. *J Appl Soc Psychol*. 1984;14:184-206.
45. US Department of Health and Human Services. *Strategies to Control Tobacco Use in the United States: A Blueprint for Public Health Action in the 1990s: Monograph No. 1*. Bethesda, MD: Public Health Service, US Dept of Health and Human Services, and National Cancer Institute, National Institutes of Health; 1991. NIH publication 92-3316.
46. Flay BR. Youth tobacco use: risks, patterns, and control. In: Slade J, Orleans CT, eds. *Nicotine Addiction: Principles and Management*. New York, NY: Oxford University Press; 1993:653-661.

Physicians Behaving Badly

THE ISSUE OF PEER RELATIONSHIPS IN MEDICINE IS BARELY mentioned in the literature. A recent MEDLINE search using key terms *professionalism*, *physician communication*, *physician relationships*, *interphysician relationships*, and *interphysician communication* yielded a wealth of writings on ethical and cultural concerns and the patient-physician relationship, but just one piece addressing clear and adequate communication *between* physicians.[1] This article addressed the problem from an educational rather than an ethical or collegial perspective, successfully testing the hypothesis that a virtual computer program could be used to enhance medical students' consultative and collaborative skills.

In the past several months numerous incidences have emerged from my own practice highlighting the problem of noncommunication in medicine.

I recently received word from a rehabilitation hospital that Anne, a patient I had followed for years, had been discharged home. The facility faxed me a brief summary of events, a set of laboratory values, and an updated medication list, and Anne, I was told, was to be seen the next day by a visiting nurse. I had never been told that Anne had been admitted for rehabilitation, nor had I been informed of the emergency department visit and hospitalization for an acute stroke that had preceded her subacute stay. During a three-week window, Anne had been treated in two separate inpatient facilities and cared for by numerous physicians, but not one of them had contacted me. I assume these physicians knew that I was Anne's primary care physician because I belatedly received an abbreviated copy of her course in the rehabilitation facility.

Anne and her 90-year-old husband lived in subsidized housing. They were retired and survived on a meager fixed income. At her last office visit Anne had been taking medications for hypertension, type 2 diabetes mellitus, and urge incontinence, which I had chosen for cost containment, dosing convenience, and effectiveness. During her hospital stay these medications were changed to more expensive and dose-inconvenient formulations about which I was never informed. Her new copayments were significantly higher than her old, and had I not reinstituted her long-standing regimen, she would have quickly reached and spent through the Medicare "doughnut hole." In addition, many of her new medications had been prescribed without regard for their effect on a 90-year-old woman whose only prior symptoms had been arthralgias, which had never affected her mobility.

In another incident, Susan, one of my most stoic and devoted patients, developed a nonhealing toe ulcer. I referred her to a vascular medicine specialist, who documented severe peripheral vascular disease then sent her for surgical evaluation. Neither consultant contacted me. Several weeks later, wholly by accident, I found Susan on our inpatient surgical service. She had been admitted for a lower-extremity bypass, had had a difficult and depressing admission, and had then been scheduled for transfer to rehabilitation. At no point during her preoperative evaluation or inpatient treatment had either of her consulting physicians updated me on her care. After discovering Susan on the vascular surgeon's inpatient service, I personally spoke with him and asked that in the future he keep me abreast of her progress. However, I was not informed of either of her following two admissions—one for a wound infection at the bypass site, and the next for a below-knee amputation when the bypass failed.

In another case, Bob, a refreshingly blunt retired engineer, had been my patient for two years when he was diagnosed with stomach cancer. I provided his general care, but his oncologist and psychiatrist were essential to our team. I called Bob at home last month just to check up on him and left a message on his answering machine. His daughter telephoned me later to tell me that he was dead! Two weeks earlier he had been admitted to a community hospital and diagnosed with multilobar pneumonia. He had opted for comfort measures and passed away the next day. Not one of Bob's hospitalists (he was apparently seen by several in succession) had contacted any of his regular physicians. As a result his family had buried him, sat shiva, and gone home, all without a timely condolence from his primary medical team.

These vignettes point to a problem of noncommunication among physicians. Most of us, given time to reflect, would consider it disrespectful to dramatically alter a patient's regimen, to operate on him, or to fill out his death certificate without first calling at least one of his primary physicians. Yet such behaviors are so common in modern practice that many of us may no longer even notice them.

These stories also reflect the difficult fact that such lack of communication has real consequences for the quality of care we deliver. Anne was prescribed medicines she could neither easily take nor afford, Susan was forced to make crucial medical decisions without the counsel of the physician she trusted and knew best, and an oversight by Bob's inpatient team left his family disappointed and his primary medical team embarrassed.

A Piece of My Mind Section Editor: Roxanne K. Young, Associate Editor.

Further, most physicians would likely agree that by inadequately collaborating with others involved in our patients' care, we miss critical pieces of their histories and thus either perform redundant workups or fail to address important medical issues. This very point was elegantly demonstrated by Moore et al,[2] who found that poor communication between inpatient and outpatient physicians at the time of hospital discharge significantly reduced the likelihood that recommended follow-up studies will be performed.

The current trend toward noncommunication has grave consequences for the future of the patient-physician relationship. Physicians are increasingly open to scrutiny and our patients are more aware than ever of our fallibility, yet our ability to deliver care *requires* our patients' confidence. When our patients sense that their physicians are not communicating appropriately with each other or, worse, undoing one another's work, their belief in our *collective* ability to care for them deteriorates. As a family of professionals, we should never let this happen.

Partial explanation for this breakdown of communication in our profession may rest with the expectations by HMOs, insurance companies, and physicians' employers to see ever-increasing numbers of patients, thus decreasing our time to communicate. However, *the fault is primarily ours* for being too inconsiderate or lazy to call, e-mail, or write our colleagues to provide information and relay our concerns about our mutual patients who have entrusted their and their loved ones' care to us. This sobering realization and knowledge that we, too, can and certainly will be patients should make us behave responsibly to our patients and colleagues. Until we do, we may be at best proficient in the science of medicine, but will never master the more difficult art.

While Anne's, Susan's, and Bob's stories are discouraging, and while in moments of contemplation we may find ourselves appalled by how poorly we treat one another, there is certainly hope for the future. Last week one of my patients was admitted to the neurology service at a local teaching hospital. When I visited him the next morning, I introduced myself to his resident and told her a bit about his previous care. In only her second year of training, this house officer had an innate appreciation of proper communication. She e-mailed me midweek to discuss my patient's progress and on the very day of his discharge sent me a summary and a list of his follow-up appointments. I complimented both the young physician and her program directors by e-mail, and they were all delighted to hear such positive feedback.

All behavior is modeled, taught, and reinforced. If we expect no more from one another than what we currently receive, our workplace milieu, our ability to deliver excellent patient care, and our professional standing will continue to slide. Until recently, I had never considered it necessary to ask hospitals for discharge summaries of my admitted patients, but I now realize that I need to request what I had previously considered an expected courtesy. I have also decided to withhold referrals to physicians who don't provide me with timely resumes of their thoughts or who don't involve me in important health care decisions involving my patients.

This problem of poor communication needs attention and correction. Perhaps responsibility should rest with medical schools' admission committees, curriculum committees, and residency program directors. Some students innately possess these interpersonal skills while for others, communication between colleagues can be taught, role modeled, and even enforced by experts. Finding those experts or rewarding them appropriately in today's medical climate may not be easy but is worth the try, considering the stakes.

Michael D. Stillman, MD
Boston, Massachusetts
michael.stillman@bmc.org

1. Sijstermans R, Jaspers MW, Bloemendaal PM, Schoonderwaldt EM. Training inter-physician communication using the Dynamic Patient Stimulator. *Int J Med Inform*. 2007;76(5-6):336-343.
2. Moore C, McGinn T, Halm E. Tying up loose ends: discharging patients with unresolved medical issues. *Arch Intern Med*. 2007;167(12):1305-1311.

Addiction

RESEARCH REPORT

doi:10.1111/j.1360-0443.2008.02245.x

The impact of alcohol use on depressive symptoms in human immunodeficiency virus-infected patients†

Lynn E. Sullivan[1], Richard Saitz[2,3], Debbie M. Cheng[2,4], Howard Libman[5], David Nunes[6] & Jeffrey H. Samet[2,7]

Yale University School of Medicine, New Haven, CT, USA,[1] Clinical Addiction Research and Education Unit, Section of General Internal Medicine, Department of Medicine, Boston Medical Center, Boston University School of Medicine, Boston, MA, USA,[2] Department of Epidemiology and Youth Alcohol Prevention Center, Boston University School of Public Health, Boston, MA, USA,[3] Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA,[4] Divisions of General Medicine and Primary Care and Infectious Diseases, Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA,[5] Section of Gastroenterology, Department of Medicine, Boston Medical Center, Boston University School of Medicine, Boston, MA, USA[6] and Department of Social and Behavioral Sciences, Boston University School of Public Health, Boston, MA, USA[7]

ABSTRACT

Aims To examine the impact of alcohol use on depressive symptoms in human immunodeficiency virus (HIV)-infected patients. **Design** Data were collected at 6-month intervals and analyzed to evaluate the association between alcohol dependence and consumption on depressive symptoms using longitudinal mixed-effects regression models controlling for specified covariates. **Measurements** The two independent variables were current alcohol dependence assessed using the Composite International Diagnostic Interview (CIDI) and past month consumption (heavy versus not heavy drinking) using a validated calendar-based method. The primary outcome was depressive symptoms as measured by the Center for Epidemiologic Studies Depression Scale (CES-D). **Participants** HIV-infected adults with current or past alcohol problems. **Findings** Alcohol dependence and heavy alcohol use were significantly associated with higher CES-D scores in unadjusted models. In adjusted analyses, the association of current alcohol dependence persisted [mean difference in CES-D was 3.49 for dependence versus non-dependence; 95% confidence interval (CI): 1.76–5.22]; however, the effect of heavy drinking was no longer statistically significant (mean difference in CES-D was 1.04 for heavy versus not heavy drinking; 95% CI: –0.24–2.32). **Conclusions** Alcohol use is associated with more depressive symptoms in HIV-infected patients with alcohol problems. This association remains significant after adjusting for potential confounders only when alcohol use meets the criteria for alcohol dependence.

Keywords Acquired immunodeficiency syndrome, alcohol drinking, alcoholism, depression, depressive disorder, HIV.

Correspondence to: Lynn E. Sullivan, Yale University School of Medicine, 367 Cedar Street, PO Box 208093, New Haven, CT 06520-8093, USA. E-mail: lynn.sullivan@yale.edu

Submitted 25 July 2007; initial review completed 9 December 2007; final version accepted 20 March 2008

INTRODUCTION

The life-time prevalence of an alcohol use disorder is higher in people living with human immunodeficiency virus/acquired immune deficiency syndrome (HIV/AIDS) (26–60%) than it is in the general population (14–24%) [1–9]. In a national sample of HIV-infected patients, 8–12% were heavy drinkers, a proportion approximately twice that of the US national average [10,11]. Depressive symptoms are also more common in HIV-infected patients, with studies reporting a life-time prevalence of depression ranging from 22% to 45% compared to 4% in the general population [12–17].

Alcohol use and depressive symptoms impact significantly upon the course of each other [18–20]. Given the high prevalence of these conditions in HIV-infected patients, they are likely to co-occur more frequently in this patient population.

Both alcohol use and depressive symptoms have a substantial impact on HIV-related behaviors and disease outcomes. Alcohol use has been associated with suboptimal utilization of medical services. For example, patients with

†This work was presented at the Annual Meeting of the Society of General Internal Medicine, 28 April 2006, Los Angeles, California.

alcohol use disorders delay seeking treatment for HIV infection [21]. Alcohol problems in HIV-infected patients are associated with poor adherence to antiretroviral therapy (ART) [8,11,22–25], worse treatment response and more rapid HIV disease progression, as evidenced by lower CD4 lymphocyte counts and higher HIV RNA [23,24], and an increase in high-risk sexual behaviors [26–30]. Similarly, depressive symptoms have a multi-faceted effect on HIV-infected people, including increased immune dysfunction, biochemical alterations and adverse effects on medication adherence [31–33]. In addition, depressive symptoms in HIV-infected patients have been associated with increased deaths, with one study revealing a mortality rate in HIV-infected women with depressive symptoms of twice that found in those without depressive symptoms [34].

In sum, both alcohol and depressive symptoms co-occur in adults and more often in those with HIV infection; both can impact HIV behaviors and outcomes. While these findings are compelling, research to date has not examined the impact of alcohol use on depressive symptoms in HIV-infected patients. Knowledge regarding how these factors relate to each other, particularly how alcohol consumption can affect depressive symptoms, is important because, if associated strongly, addressing alcohol use could have substantial impact on both depressive symptoms and HIV behaviors and outcomes. Therefore, the purpose of this study was to determine if current alcohol dependence and alcohol consumption affect depressive symptoms in people with HIV infection. We studied this association in a prospective cohort of HIV-infected patients with past or current alcohol problems.

METHODS

Study design

We conducted an analysis of data from a prospectively followed cohort of HIV-infected subjects enrolled in the HIV-LIVE (HIV–Longitudinal Interrelationships of Viruses and Ethanol) study between August 2001 and July 2003. Data on alcohol use behaviors and depressive symptoms were collected prospectively every 6 months by trained interviewers using standardized instruments. In-person assessment interviews and laboratory tests were scheduled every 6 months for a total of 36 months of follow-up. We examined the association between alcohol dependence and consumption with depressive symptoms in this cohort of HIV-infected adults with current or past alcohol problems.

Eligibility criteria included the following: HIV infection documented by HIV antibody test by enzyme-linked immunosorbent assay (ELISA) with Western blot confirmation; current or past alcohol problems supported either by two or more positive responses to the CAGE (Cut-down, Annoyed, Guilty, Eye-opener) alcohol screening questionnaire [35,36] or by a study physician investigator's clinical determination of alcohol abuse or dependence; the ability to speak English or Spanish; identification of a contact person who knew the subject's whereabouts; a score of greater than 20 on the Folstein Mini-Mental State Examination [37,38]; and a trained interviewer assessment that the subject was capable of giving informed consent and answering the interview questions. Subjects were recruited from several different sources, including: (1) a previous cohort study of people with HIV and alcohol problems ($n = 154$, 38%) [25]; (2) the Diagnostic Evaluation Unit (DEU), an intake clinic for HIV-infected patients at Boston Medical Center (BMC) ($n = 87$, 22%) [39]; (3) the HIV Primary Care and Specialty Clinics at Beth Israel Deaconess Medical Center (BIDMC) ($n = 31$, 8%); and (4) additional health care centers, homeless shelters, drug treatment programs, subject referrals and flyers ($n = 128$, 32%). The Institutional Review Boards of Boston Medical Center and Beth Israel Deaconess Medical Center approved this study. Additional privacy protection was secured by the issuance of a Certificate of Confidentiality by the Department of Health and Human Services to protect subjects from release of their research data even under a court order or subpoena.

Measurements

The primary outcome variable was the Center for Epidemiologic Studies, Depression Scale (CES-D) [40] score. The CES-D is a 20-item self-report questionnaire used to assess the presence and severity of depressive symptoms, with scores ranging from 0 to 60.

One of the two independent variables of interest was current alcohol dependence, defined as a diagnosis of alcohol dependence by meeting diagnostic criteria in the past 6 months and assessed using the reference standard Composite International Diagnostic Interview (CIDI) [41]. The second independent variable was past-month alcohol consumption [heavy drinking (more than four drinks on 1 day or more than 14 drinks per week on average for men; more than three or more than seven drinks, respectively, for women) versus not heavy drinking (none or moderate) amounts] as assessed using a validated calendar-based method [42]. We stratified the cohort further into abstinent (no alcohol consumption), moderate (any alcohol consumption but not heavy drinking), heavy drinking and very heavy drinking (>4 separate days of more than four drinks on 1 day for men; >4 separate days of more than three drinks on 1 day for women).

Other specific subject characteristics assessed included age, sex, race/ethnicity, homelessness (defined

as having spent at least one night either on the street or in a shelter in the 6 months before the interview) [43] and hepatitis C (HCV) antibody status. We also assessed for the presence of other medical comorbidities (using the Katz comorbidity scale) [44]. Medical diagnoses were read to the participants, and they indicated whether they had been hospitalized for these reasons in the previous 6 months. Data were collected on antiretroviral medication use and adherence, past month heroin and cocaine use, CD4 cell counts and HIV log RNA measurements as measured using branched-chain DNA techniques [45]. Adherence to ART was determined using the AIDS Clinical Trials Group Questionnaire for Adherence to Anti-Retroviral Medications [46]. Subjects who reported being less than 100% adherent during the previous 3 days were considered not adherent.

Participants

The baseline characteristics of the 400 subjects enrolled in the HIV-LIVE study included a mean age of 43 years [standard deviation (SD) = 7.4, range = 21–71], 75% were men, 41% were black, 33% were white, 19% were Hispanic and 7% were other races/ethnicities. Twenty-five per cent of the sample were homeless, 59% (232 of 396) were hepatitis C antibody positive, 64% reported current illicit drug use. Ten per cent met criteria for alcohol dependence, 31% reported heavy drinking, 11% reported moderate drinking and 58% reported no alcohol consumption. The mean CES-D score was 22 (SD = 12.9, range = 0–56); 46% reported adherence to antiretroviral medication, 16% reported non-adherence and 38% reported not being on any antiretroviral medications. The mean CD4 count was 455 cells/mm^3 (SD = 299, range = 8–1809) and the mean HIV log RNA was 2.98 copies/ml (SD = 1.35, range = 1.5–5.88). The median number of study visits per subject was four (interquartile range three to six visits) and the median time between baseline and last follow-up visit (in months) was 23.6 months (interquartile range 16.7–30.3 months).

Data analyses

Analyses of demographic and clinical characteristics at baseline included descriptive statistics. Separate linear mixed effects multiple regression models [47] were used to examine the association between both alcohol dependence and current alcohol consumption and depressive symptoms. Alcohol dependence and consumption were modeled as time-varying variables. The models controlled for the following covariates: gender, age, race/ethnicity (black versus non-black), homelessness (yes versus no), HCV antibody status (positive versus negative), the Katz comorbidity scale, any past month illicit drug use (yes versus no), antiretroviral medication use and adherence (three-category variable: not on ART, on ART but not adherent, adherent ART), CD4 cell counts, HIV log RNA measurements and time since study enrollment (months). Age, Katz comorbidity scale, CD4 cell count and HIV log RNA were included in regression models as continuous variables. All other covariates were included as categorical variables using dummy variables. ART medication status, illicit drug use, CD4 cell count, HIV log RNA and time since study enrollment were included as time-varying covariates. All other covariates were taken from the baseline assessment. The linear mixed models included a random intercept and random slope for each subject to account for the correlation from including multiple observations from the same subject. Models were fitted using an unstructured variance covariance matrix and empirical standard errors were reported for all analyses. All HIV-LIVE subjects were included in the current analyses. To minimize the potential for colinearity, we assessed correlation between each pair of independent variables and verified that no pair of variables included in the same regression model was highly correlated (i.e. >0.40). All analyses used two-tailed tests of significance and were performed using SAS software (version 8.2; SAS Institute, Cary, NC, USA). *P*-values <0.05 were considered statistically significant.

RESULTS

Using longitudinal regression models, we found that current alcohol dependence was associated with higher CES-D scores in unadjusted models (24 versus 21; $P < 0.0001$) and that the association persisted in adjusted analyses (24 versus 21; $P < 0.0001$) (Table 1). The impact of alcohol dependence and the specified covariates on depressive symptoms are displayed in Table 2. Subjects with alcohol dependence, who were HCV antibody-positive or currently using illicit drugs exhibited significantly higher CES-D scores (all *P*-values <0.01). Subjects who were black or who were adherent to their ART exhibited significantly lower CES-D scores (all *P*-values <0.05). Additional exploratory analyses controlling for marital/partner status produced similar findings [mean difference in CES-D was 3.44 for dependence versus non-dependence; 95% confidence interval (CI): 1.73–5.16; mean difference in CES-D was 1.08 for heavy versus not heavy drinking; 95% CI: –0.20–2.35].

In contrast, while the unadjusted mean CES-D scores were significantly higher for heavy drinkers compared with those who were not (23 versus 21; $P = 0.005$) (Table 3), in the adjusted analysis controlling for the specified covariates, the difference in the mean CES-D

Table 1 The association between current alcohol dependence and depressive symptoms.

	Currently alcohol-dependent	*Not alcohol-dependent*	*Mean difference (95% CI)*	*P-value*
Unadjusted mean depressive symptoms (SE)*	24 (0.93)	21 (0.57)	3.54 (1.89, 5.19)	<0.0001
Adjusted† mean depressive symptoms (SE)‡	24 (1.05)	21 (0.77)	3.49 (1.76, 5.22)	<0.0001

*Analyses based on 400 subjects contributing 1721 observations. †The adjusted mean scores control for gender, age, race/ethnicity (black versus non-black), homelessness, hepatitis C virus antibody status, the Katz comorbidity scale, past month illicit drug use, antiretroviral therapy medication use and adherence, CD4 cell counts, human immunodeficiency virus log RNA measurements, and time in months since study enrollment. ‡Analyses based on 391 subjects contributing 1509 observations. CI: confidence interval; SE: standard deviation.

Table 2 Multivariable analysis of the association between current alcohol dependence and covariates with depressive symptoms.

Characteristic	*Adjusted mean difference in depressive symptoms (SE)*	*95% confidence interval*	*P-value*
Currently alcohol-dependent	3.49 (0.88)	1.76,5.22	<0.0001
Female	2.36 (1.34)	–0.28,5.00	0.08
Age	–0.03 (0.08)	–0.18,0.13	0.73
Black race	–2.52 (1.14)	–4.76,–0.27	0.03
Homelessness	1.43 (1.29)	–1.10,3.96	0.27
HCV antibody positive	4.60 (1.17)	2.31,6.89	<0.0001
Katz medical comorbidity	0.05 (0.18)	–0.30,0.40	0.78
Current illicit drug use	1.71 (0.63)	0.48,2.94	0.006
HIV log RNA	0.41 (0.28)	–0.15,0.96	0.15
CD4 cell count	–0.002 (0.001)	–0.0046,0.0006	0.13
Adherent to highly active antiretroviral therapy versus not on medications	–1.76 (0.88)	–3.48,–0.04	0.04
Not adherent to highly active antiretroviral therapy versus not on medications	–1.22 (0.91)	–3.00,0.56	0.18

HIV: human immunodeficiency virus; SE: standard error.

Table 3 The association between current heavy drinking and depressive symptoms.

	Current heavy drinking	*Not heavy drinking*	*Mean difference (95% CI)*	*P-value*
Unadjusted mean depressive symptoms (SE)*	23 (0.75)	21 (0.58)	1.76 (0.53, 2.98)	0.005
Adjusted† mean depressive symptoms (SE)‡	22 (0.90)	21 (0.77)	1.04 (–0.24, 2.32)	0.11

*Analyses based on 400 subjects contributing to 1726 observations. †The adjusted mean scores control for gender, age, race/ethnicity (black versus non-black), homelessness, hepatitis C virus antibody status, the Katz comorbidity scale, past month illicit drug use, antiretroviral therapy medication use and adherence, CD4 cell counts, human immunodeficiency virus log RNA measurements, and time in months since study enrollment. ‡Analyses based on 391 subjects contributing to 1514 observations. CI: confidence interval; SE: standard error.

scores was attenuated and no longer significant (22 versus 21; $P = 0.11$).

When this relationship was assessed further using the four drinking categories, depressive symptoms appeared to increase as drinking levels increase [adjusted mean CES-D scores (SD): non-drinkers 21 (0.81); moderate drinkers 21 (1.01); heavy drinkers 22 (0.93); very heavy drinkers 23 (1.15)]; however, the differences were not statistically significant.

DISCUSSION

As both unhealthy alcohol use and depressive symptoms are common among HIV-infected individuals, understanding the relationship between these two comorbidities has significance for optimal patient care. Our study found that current alcohol dependence was associated independently with more depressive symptoms in HIV-infected patients with current or past alcohol problems.

We did not detect a statistically significant association between heavy drinking and depressive symptoms. Finally, our results substantiate previous work that demonstrated additional factors, such as being HCV antibody-positive or currently using illicit drugs being associated with depressive symptoms among HIV-infected patients with alcohol problems [48]. Our study also found that those who were black or adherent to their antiretroviral medications had significantly lower CES-D scores. This relationship between race and depression in patients with alcohol dependence has also been found in a large national sample [49]. Similarly, the relationship we found between adherence to antiretroviral medications and depression has been demonstrated by a number of studies, although the directionality of the relationship (i.e. change in adherence leading to change in depressive symptoms versus change in depressive symptoms leading to change in adherence) is not clear [25,50,51].

The relationship between substance abuse and mental health was highlighted recently by the Institute of Medicine (IOM) in its report entitled *Improving the Quality of Health Care for Mental and Substance-Use Conditions*: Quality Chasm Series [52]. This report has as a fundamental principle that the physical health of the patient is linked tightly to his/her substance use status and mental health. The current results demonstrate that, in addition, these important mental health and substance use factors relate not only to physical health such as HIV disease, but also relate to each other. Depressive symptoms are not independent of the use and consequences of one's alcohol use. Our study highlights the important relationship between alcohol use and depressive symptoms in patients with HIV disease. The importance of depressive symptoms among HIV-infected people has been noted previously, perhaps most impressively in a previous study in HIV-infected women in which depressive symptoms, as measured by the CES-D scale, were associated with HIV disease progression and death [34]. Therefore, any worsening of depressive symptoms is of major concern and any factor that might ameliorate depressive symptoms is of great interest.

A recent systematic review of the literature focusing specifically upon the impact of alcohol problems in individuals with major depression found that alcohol problems are common in people with depression and that they are associated with a number of adverse clinical and health care outcomes [18]. While the majority of these data focused upon the effect of alcohol dependence on depression outcomes and revealed findings similar to our study, they did not provide data on the impact of lower levels of alcohol consumption, such as heavy or problem drinking, on depressive symptoms. A subsequent study by Alati *et al.* did find a linear relationship between alcohol consumption and depressive and anxiety symptoms at certain stages in the lives of their subjects, but also concluded that this relationship was dependent upon the gender and age of the subjects [53]. Of note, we are unaware of any similar studies that have examined the impact of alcohol dependence or use on depression outcomes specifically in HIV-infected patients.

Our study has several limitations. Although we did not detect an association between current heavy drinking and depressive symptoms in this cohort, it is important to note that our study cohort had a significant burden of depressive symptoms at baseline and therefore may be masking some of the impact that varying levels of alcohol had on these symptoms. Furthermore, while the CES-D is recognized as a well-validated scale that has been used extensively in medically ill patient populations to measure depressive symptoms, it is not a diagnostic instrument for depression. A recent study examining the association between alcohol and depression highlighted that there are certain factors such as gender and methods of measuring these two conditions that must be taken into consideration when examining the alcohol–depression relationship [54]. Also, the current study does not provide information on the use of antidepressant therapy in these patients which, if present, may have mitigated the impact of alcohol use. Finally, this study was potentially underpowered to detect effects of the observed magnitude for heavy drinking. Although the observed one-unit increase in depressive symptoms in subjects with heavy alcohol use compared to those without heavy alcohol use may be clinically important, it was not statistically significant and the resulting confidence intervals were wide. Thus, we are also unable to conclude that no association exists. A larger cohort may be necessary to provide definitive conclusions on the effect of heavy drinking.

In summary, we found in our study that HIV-infected patients with alcohol dependence had significantly more depressive symptoms. These findings highlight the impact on depressive symptoms of the severe end of the spectrum of alcohol use disorders, alcohol dependence. Although an increase in depressive symptoms was noted among those with heavy drinking (and across increasing levels of alcohol use), it was small and not statistically significant. In addition, there are specific potentially modifiable factors, including current illicit drug use and antiretroviral medication adherence, associated with this relationship.

The findings of our study are important, as they provide information about an HIV-infected population *per se* and present data regarding the impact of both alcohol dependence and heavy drinking on depressive symptoms. Future investigations should evaluate systematically the effect that varying levels of alcohol use have on depressive symptoms in a larger cohort of HIV-infected patients. In

addition, they should explore the impact of alcohol use on the effectiveness of pharmacological and/or psychotherapeutic treatment for depression in this population. Finally, future work should focus upon how different interventions treating alcohol problems ultimately affect various aspects of HIV disease.

Acknowledgments

Support for this study came from the following grants from the National Institute on Alcohol Abuse and Alcoholism (NIAAA) of the NIH: R01-AA13216 (Clinical Impact of HCV and Alcohol in HIV-Infected Persons), R01-AA11785 (Medication Adherence in Alcohol Abusing HIV Patients); R01-AA10870 (Enhanced Linkage of Alcohol Abusers to Primary Care); and K24-AA015674 (Impact of Alcohol Use on HIV Infection—in US and Russia) (J. S.). This study was also supported, in part, by the National Institute on Drug Abuse (NIDA) Physician Scientist Award (K12 DA00167) and the Robert Wood Johnson Physician Faculty Scholars Program (L. S.). This research was conducted in part in the General Clinical Research Center at Boston University School of Medicine, USPHS Grant M01 RR00533, and the Beth Israel Deaconess General Clinical Research Center, USPHS Grant M01 RR01032.

References

1. Williams J. B., Rabkin J. G., Remien R. H., Gorman J. M., Ehrhardt A. A. Multidisciplinary baseline assessment of homosexual men with and without human immunodeficiency virus infection. II. Standardized clinical assessment of current and lifetime psychopathology. *Arch Gen Psychiatry* 1991; **48**: 124–30.
2. Waller S. J., Lyons J. S., Costantini-Ferrando M. F. Impact of comorbid affective and alcohol use disorders on suicidal ideation and attempts. *J Clin Psychology* 1999; **55**: 585–95.
3. Samet J., Phillips S. J., Horton N. J., Traphagen E. T., Freedberg K. A. Detecting alcohol problems in HIV-infected patients: use of the CAGE questionnaire. *AIDS Res Hum Retroviruses* 2004; **20**: 151–5.
4. Regier D. A., Farmer M. E., Rae D. S., Locke B. Z., Keith S. J., Judd L. L. *et al*. Comorbidity of mental disorders with alcohol and other drug abuse. Results from the Epidemiologic Catchment Area (ECA) study. *JAMA* 1990; **264**: 2511–18.
5. Petry N. M. Alcohol use in HIV patients: what we don't know may hurt us. *Int J STD AIDS* 1999; **10**: 561–70.
6. Lefevre F., O'Leary B., Moran M., Mossar M., Yarnold P. R., Morton G. J. *et al*. Alcohol consumption among HIV-infected patients. *J Gen Intern Med* 1995; **10**: 458–60.
7. Kessler R. C., McGonagle K. A., Zhao S., Nelson C. B., Hughes M., Eshleman 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.
8. Cook R. L., Sereika S. M., Hunt S. C., Woodward W. C., Erlen J. A., Conigliaro J. Problem drinking and medication adherence among persons with HIV infection. *J Gen Intern Med* 2001; **16**: 83–8.
9. Dew M. A., Becker J. T., Sanchez J., Caldararo R. Prevalence and predictors of depressive, anxiety and substance use disorders in HIV-infected and uninfected men: a longitudinal evaluation. *Psychol Med* 1997; **27**: 395–409.
10. Burnam M. A., Bing E. G., Morton S. C., Sherbourne C., Fleishman J. A., London A. S. *et al*. Use of mental health and substance abuse treatment services among adults with HIV in the United States. *Arch Gen Psychiatry* 2001; **58**: 729–36.
11. Galvan F. H., Bing E. G., Fleishman J. A., London A. S., Caetano R., Burnam M. A. *et al*. The prevalence of alcohol consumption and heavy drinking among people with HIV in the United States: results from the HIV Cost and Services Utilization study. *J Stud Alcohol* 2002; **63**: 179–86.
12. Brown G. R., Rundell J. R., McManis S. E., Kendall S. N., Zachary R., Temoshok L. Prevalence of psychiatric disorders in early stages of HIV infection. *Psychosom Med* 1992; **54**: 588–601.
13. Brown G. R., Rundell J. R. A prospective study of psychiatric aspects of early HIV disease in women. *Gen Hosp Psychiatry* 1993; **15**: 139–47.
14. McDaniel J. S., Fowlie E., Summerville M. B., Farber E. W., Cohen-Cole S. A. An assessment of rates of psychiatric morbidity and functioning in HIV disease. *Gen Hosp Psychiatry* 1995; **17**: 346–52.
15. Perkins D. O., Stern R. A., Golden R. N., Murphy C., Naftolowitz D., Evans D. L. Mood disorders in HIV infection: prevalence and risk factors in a nonepicenter of the AIDS epidemic. *Am J Psychiatry* 1994; **151**: 233–36.
16. Kelly B., Raphael B., Judd F., Perdices M., Kernutt G., Burrows G. D. *et al*. Psychiatric disorder in HIV infection. *Aust NZ J Psychiatry* 1998; **32**: 441–53.
17. Narrow W. E., Rae D. S., Robins L. N., Regier D. A. Revised prevalence estimates of mental disorders in the United States: using a clinical significance criterion to reconcile 2 surveys' estimates. *Arch Gen Psychiatry* 2002; **59**: 115–23.
18. Sullivan L. E., Fiellin D. A., O'Connor P. G. The prevalence and impact of alcohol problems in major depression: a systematic review. *Am J Med* 2005; **118**: 330–41.
19. Goldstein B. I., Levitt A. J. Is current alcohol consumption associated with increased lifetime prevalence of major depression and suicidality? Results from a pilot community survey. *Compr Psychiatry* 2006; **47**: 330–3.
20. Saitz R. Clinical practice. Unhealthy alcohol use. *N Engl J Med* 2005; **352**: 596–607.
21. Samet J. H., Freedberg K., Stein M. D., Lewis R., Savetsky J., Sullivan L. *et al*. Trillion virion delay: time from testing positive for HIV to presentation for primary care. *Arch Intern Med* 1998; **158**: 734–40.
22. Fabris P., Tositti G., Manfrin V., Giordani M. T., Vaglia A., Cattelan A. M. *et al*. Does alcohol intake affect highly active antiretroviral therapy (HAART) response in HIV-positive patients? *J Acquir Immune Defic Syndr* 2000; **25**: 92–3.
23. Lucas G. M. Longitudinal assessment of the effects of drug and alcohol abuse on HIV-1 treatment outcomes in an urban clinic. *AIDS* 2002; **16**: 767–74.
24. Samet J. H., Horton N. J., Traphagen E. T., Lyon S. M., Freedberg K. A. Alcohol consumption and HIV disease progression: are they related? *Alcohol Clin Exp Res* 2003; **27**: 862–7.
25. Samet J. H., Horton N. J., Meli S., Freedberg K. A., Palepu A. Alcohol consumption and antiretroviral adherence among HIV-infected persons with alcohol problems. *Alcohol Clin Exp Res* 2004; **28**: 572–7.
26. Stein M., Herman D. S., Trisvan E., Pirraglia P., Engler P.,

Anderson B. J. Alcohol use and sexual risk behavior among human immunodeficiency virus-positive persons. *Alcohol Clin Exp Res* 2005; **29**: 837–43.
27. Scheidt D. M., Windle M. A comparison of alcohol typologies using HIV risk behaviors among alcoholic inpatients. *Psychol Addict Behav* 1997; **11**: 3–17.
28. McMahon R. C., Malow R. M., Jennings T. E., Gomez C. J. Effects of a cognitive-behavioral HIV prevention intervention among HIV negative male substance abusers in VA residential treatment. *AIDS Educ Prev* 2001; **13**: 91–107.
29. Woods W. J., Lindan C. P., Hudes E. S., Boscarino J. A., Clark W. W., Avins A. L. HIV infection and risk behaviors in two cross-sectional surveys of heterosexuals in alcoholism treatment. *J Stud Alcohol* 2000; **61**: 262–66.
30. Ehrenstein V., Horton N. J., Samet J. H. Nonconsistent condom use among HIV-infected patients with alcohol problems. *Drug Alcohol Depend* 2004; **73**: 159–66.
31. Turner B. J., Laine C., Cosler L., Hauck W. W. Relationship of gender, depression, and health care delivery with antiretroviral adherence in HIV-infected drug users. *J Gen Intern Med* 2003; **18**: 248–57.
32. Cruess D. G., Petitto J. M., Leserman J., Douglas S. D., Gettes D. R., Ten Have T. R. *et al*. Depression and HIV infection: impact on immune function and disease progression. *CNS Spectr* 2003; **8**: 52–8.
33. Leserman J. HIV disease progression: depression, stress, and possible mechanisms. *Biol Psychiatry* 2003; **54**: 295–306.
34. Ickovics J. R., Hamburger M. E., Vlahov D. Mortality, CD4 cell count decline, and depressive symptoms among HIV-seropositive women: longitudinal analysis for the HIV Epidemiology Research study. *JAMA* 2001; 285.
35. Ewing J. A. Detecting alcoholism: the CAGE questionnaire. *JAMA* 1984; **252**: 1905–07.
36. Fiellin D. A., Reid M. C., O'Connor P. G. Screening for alcohol problems in primary care: a systematic review. *Arch Intern Med* 2000; **160**: 1977–89.
37. Folstein M. F., Folstein S. E., McHugh P. R. 'Mini-mental state'. A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975; **12**: 189–98.
38. Smith K. L., Horton N. J., Saitz R., Samet J. H. Use of mini-mental state examination in recruitment for substance abuse research studies. *Drug Alcohol Depend* 2006; **82**: 231–7.
39. Samet J. H., Libman H., LaBelle C., Steger K., Lewis R., Craven D. E. *et al*. A model clinic for the initial evaluation and establishment of primary care for persons infected with human immunodeficiency virus. *Arch Intern Med* 1995; **155**: 1629–33.
40. Radloff L. S. The CES-D scale: a self-report depression scale for research in the general population. *Appl Psychol Meas* 1977; **1**: 385–401.
41. Robins L. N., Wing J., Wittchen H. U. 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.
42. Sobell L. C., Sobell M. B. Alcohol timeline follow-back (TLFB). In: American Psychiatric Association. *Handbook of Psychiatric Measures*. Washington DC: American Psychiatric Association; 1996; p. 477–79.
43. Kertesz S. G., Horton N. J., Friedmann P. D., Saitz R., Samet J. H. Slowing the revolving door: stabilization programs reduce homeless persons' substance use after detoxification. *J Subst Abuse Treat* 2003; **24**: 197–207.
44. Katz J. N., Chang L. C., Sangha O., Fossel A. H., Bates D. W. Can comorbidity be measured by questionnaire rather than medical record review? *Med Care* 1996; **34**: 73–84.
45. Pachl C., Todd J. A., Kern D. G., Sheridan P. J., Fong S. J., Stempien M. *et al*. Rapid and precise quantification of HIV-1 RNA in plasma using a branched DNA signal amplification assay. *J Acquir Immune Defic Syndr* 1995; 8.
46. Chesney M. A., Ickovics J. R., Chambers D. B., Gifford A. L., Neidig J., Zwickl B. *et al*., Patient Care Committee and Adherence Working Group of the Outcomes Committee of the Adult AIDS Clinical Trials Group (AACTG). Self-reported adherence to antiretroviral medications among participants in HIV clinical trials: the AACTG adherence instruments. *AIDS Care* 2000; **12**: 255–66.
47. Laird N. M., Ware J. H. Random-effects models for longitudinal data. *Biometrics* 1982; **38**: 963–74.
48. Libman H., Saitz R., Nunes D., Cheng D. M., Richardson J. M., Vidaver J. *et al*. Hepatitis C infection is associated with depressive symptoms in HIV-infected adults with alcohol problems. *Am J Gastroenterol* 2006; **101**: 1804–10.
49. Grant B. F., Harford TC. Comorbidity between DSM-IV alcohol use disorders and major depression: results of a national survey. *Drug Alcohol Depend* 1995; **39**: 197–206.
50. Chan K. S., Orlando M., Joyce G., Gifford A. L., Burnam M. A., Tucker J. S. *et al*. Combination antiretroviral therapy and improvements in mental health: results from a nationally representative sample of persons undergoing care for HIV in the United States. *J Acquir Immune Defic Syndr* 2003; **33**: 104–11.
51. Tucker J. S., Burnam M. A., Sherbourne C. D., Kung F.-Y., Gifford A. L. Substance use and mental health correlates of nonadherence to antiretroviral medications in a sample of patients with human immunodeficiency virus infection. *Am J Med* 2003; **114**: 573–80.
52. The Institute of Medicine Committee on Crossing the Quality Chasm. *Improving the Quality of Health Care for Mental and Substance-Use Conditions: Quality Chasm Series*. Adaptation to Mental Health and Addiction Disorders. Washington DC: National Academies Press; 2006.
53. Alati R., Lawlor D. A., Najman J. M., Williams G. M., Bor W., O'Callaghan M. Is there really a 'J-shaped' curve in the association between alcohol consumption and symptoms of depression and anxiety? Findings from the Mater-University Study of Pregnancy and its outcomes. *Addiction* 2005; **100**: 643–51.
54. Graham K., Massak A., Demers A., Rehm J. Does the association between alcohol consumption and depression depend on how they are measured? *Alcohol Clin Exp. Res* 2007;**31**: 78–88.

Dig Dis Sci (2008) 53:809–814
DOI 10.1007/s10620-007-9926-x

ORIGINAL PAPER

Treatment Eligibility and Outcomes in Elderly Patients with Chronic Hepatitis C: Results from the VA HCV-001 Study

Judith I. Tsui · Sue Currie · Hui Shen · Edmund J. Bini · Norbert Brau · Teresa L. Wright · For the VA HCV-001 Study group

Received: 8 January 2007 / Accepted: 9 July 2007 / Published online: 1 September 2007

Abstract *Objectives* We undertook this study to determine if treatment candidacy and outcomes were similar between elderly and non-elderly patients. *Methods* This was a prospective cohort study that screened 4,025 patients with chronic hepatitis C for HCV antiviral treatment at 24 Veterans Affairs Medical Centers throughout the country. We used multivariable logistic regression to determine whether there was an independent association between being elderly (age > 60 vs. $\leq$ 60) and (1) being considered a treatment candidate by clinician, and (2) achieving sustained virologic response if treated. *Results* 364 of the 4,025 patients (9%) were over the age of 60. Only 25% of patients over the age of 60 were considered to be treatment candidates by the evaluating clinician, and only 10% were started on treatment. After adjustment for potential confounders, older age remained associated with a lower likelihood of being considered a treatment candidate (adjusted OR = 0.43; 95% CI: 0.30–0.61). Although based on a small sample of elderly treated patients (n = 35), being elderly did not appear to be associated with a lower likelihood of achieving SVR (adjusted OR = 1.54; 95% CI: 0.46–5.15). *Conclusion* Among veterans over the age of 60 with chronic hepatitis C who are referred for treatment, relatively few are considered treatment candidates and an even smaller number are ultimately treated. After adjusting for co-morbidities, age remains a strong predictor of not being a treatment candidate. In contrast, older age does not seem to adversely affect treatment outcomes and side effects.

Keywords Hepatitis C Virus · Elderly

J. I. Tsui (✉) · S. Currie
San Francisco VAMC, 4150 Clement St, San Francisco, CA 94121, USA
e-mail: Judith.Tsui@ucsf.edu

J. I. Tsui · H. Shen
University of California, San Francisco, CA, USA

E. J. Bini
New York City VAMC, New York, NY, USA

N. Brau
Bronx VAMC, Bronx, NY, USA

T. L. Wright
Roche Molecular Diagnostics, Pleasanton, CA, USA

Introduction

According to national surveys, approximately 1% of Americans currently over the age of 60 years old have evidence of being infected with hepatitis C virus (HCV) [1]. Although the overall incidence of HCV continues to decrease among young persons, the most prevalent group (4.3%) of infected persons are now in the 40–49 age range and are fast approaching their fifth and sixth decades of life. Little attention, however, has been paid to the management of older patients with chronic hepatitis C and current guidelines do not specifically refer to how age should impact the evaluation and treatment of HCV [2, 3]. Given the slow rate of fibrosis associated with HCV [4, 5], treatment for HCV with interferon/ribavirin may not be indicated for many older adults in whom life expectancy is limited. However, research suggests that the rate of liver fibrosis depends on numerous factors, including duration of infection and age at infection [4–6], so older adults may also be at risk for accelerated complications. Evidence that HCV is increasingly becoming an issue for older adults is seen by the increase in complications from HCV in that age

group. The average age for liver transplant has been steadily increasing such that currently at least 20% of liver transplant recipients are older than 60 years of age [7]. Likewise, the incidence of hepatocellular cancer has been increasing steadily, with the peak incidence observed among adults 75–79 years of age [8]. Therefore, for some elderly patients without substantial co-morbidities, chronic HCV may still be an important health issue. This purpose of this study was to determine if being elderly impacts treatment candidacy and outcomes among patients referred to specialty clinics for management of their chronic hepatitis C.

Methods

Study population and design

This prospective cohort study followed patients who were recruited from gastroenterology, hepatology, and infectious disease clinics at 24 Veterans Affairs (VA) Medical Centers throughout the United States between December 1999 and December 2000. Details on study sample and data collection have been described in full previously [9, 10]. Briefly, patients were eligible for this study if they were a US veteran receiving care at one of the participating study sites, were greater than 18 years of age, had a positive HCV antibody test (Ortho HCV ELISA version 3.0; Ortho-Clinical Diagnostics, Inc., Raritan, NJ), and were under consideration for HCV treatment with interferon alpha-2b and ribavirin. Patients were excluded if they had undetectable HCV RNA by polymerase-chain-reaction testing (COBAS Amplicor HCV Monitor Test, Roche Diagnostics, Branchburg, NJ). All persons provided written informed consent to participate, and the study was approved by the local institutional review board at each medical center.

Study variables

The main outcomes of interest were treatment candidacy and treatment outcomes. Treatment candidacy by clinician was determined by asking the treating clinician at each medical center "Based on your judgment, is the patient an interferon and ribavirin treatment candidate?". Treatment candidacy criteria were based on the VA's HCV treatment recommendations at the time of the study [11]. These did not specifically include a threshold for withholding treatment based on age. The main treatment outcome of interest was sustained virologic response (SVR), or absence of detectable HCV RNA (<100 copies/ml) measured at six months after HCV antiviral therapy (interferon alpha-2b and ribavirin) completion. Secondary analyses included the evaluation of treatment acceptance, initiation, side effects and discontinuation. Our main predictor of interest was elderly status, which we defined as age greater than 60 years. Additional covariates that were used for the descriptive and multivariate analysis were: sex, race/ethnic group (non-Hispanic white, non-Hispanic black, Hispanic/Latino, and other), education (high school education or less), income (less than $10,000/year), prior history of injection drug use (IDU), current substance abuse, drinking habits (none/<3/3–6/>6 drinks per day), psychiatric disease, HIV status, cardiac disease, medical co-morbidities in general, inability to remain compliant (in the judgement of the clinician), body mass index (BMI), HCV RNA genotype and viral load, ALT level, platelet count, and liver biopsy results.

Statistical analysis

Descriptive analysis of elderly and non-elderly patients was performed, using the chi-square test to compare proportions of categorical variables. Univariate analysis evaluating the proportion of treatment candidacy and outcomes in elderly and non-elderly patients was also conducted using chi-square testing. We used logistic regression in order to assess the effects of being elderly on 1) being considered a treatment candidate by clinician and 2) likelihood of attaining SVR. Odds ratios and 95% confidence intervals were calculated for the adjusted and unadjusted association between being elderly and each outcome. The variables which we chose to adjust for in multivariable models were determined in advance based on an a priori hypothesis that they could act as confounders (both positive and negative), and results from our univariate analysis (i.e., P-value < 0.05). For the logistic model for treatment candidacy, we adjusted for sex, race/ethnicity, education, income, HIV status, abnormal ALT, history of injection drug use, current substance abuse, alcohol use, psychiatric co-morbidities, medical co-morbidities, cardiac disease and non-compliance. For the model with SVR as outcome, we adjusted for sex, race, genotype (1 vs. non-1), high viral load (>1 million copies/ml), BMI and stage of fibrosis. All analyses were conducted using SAS software (SAS Institute, Cary, NC) and a two-tailed P-value of <0.05 was considered significant for all hypothesis testing.

Results

Of the study cohort of 4,025 patients who were referred for hepatitis C treatment, 364 (9%) were older than 60 years of age. Elderly patients with chronic HCV had less education; however, there was no significant difference in income,

race/ethnic groups and sex between elderly and non-elderly patients (Table 1). Elderly patients were less likely to have a history of injection drug use, and were less likely to have recently used illicit drugs. They were less likely to drink heavily (>6 drinks a day); however, there was no significant difference in the proportion reporting recent abstinence at baseline, which was low in both groups (14 vs. 16%). Elderly patients with HCV were less likely to have psychiatric co-morbidities, but were more likely to have medical co-morbidities (with the exception of HIV). Elderly patients were less likely to have abnormal ALT, but there was no significant difference in the proportion with low platelet counts. Although only a subset of patients received genotype and viral load testing, there did not appear to be any significant differences between elderly and non-elderly patients: 68 vs. 67% (chi-square *P*-value = 0.86) were genotype 1, and 48 vs. 60% (chi-square *P*-value = 0.19) had a high viral load (>1 million copies/ml). Although elderly patients were less likely to have received liver biopsies, the prevalence with advanced liver disease (greater than stage 2 fibrosis) was nearly identical among elderly and non-elderly patients (68.4 vs. 67.4%).

Elderly patients were significantly less likely to be considered a treatment candidate by the evaluating provider, compared to non-elderly patients (25 vs. 42%) (Table 2). Likewise, according to guidelines, elderly patients were also less likely to qualify for treatment (16% vs. 26%). However, like the non-elderly, elderly patients were more often considered treatment candidates by the clinician than guidelines would indicate. Using logistic regression, we observed that elderly patients who were evaluated for their chronic HCV were significantly less likely to be considered a treatment candidate (aOR 0.43; 95% CI: 0.30–0.61), even after adjusting for sex, race, education, income, HIV status, abnormal ALT, substance abuse, alcohol use, psychiatric co-morbidities, medical co-morbidities, cardiac disease and non-compliance (Table 3).

Fewer elderly patients who were considered treatment candidates agreed to be treated compared to non-elderly (63% vs. 77%). Only 10% of elderly patients who were referred for evaluation for treatment for chronic hepatitis C initiated treatment compared to 20% of non-elderly patients. There was no significant difference in the rate of side-effect and early discontinuation of treatment between elderly and non-elderly patients (Table 4): up to a quarter

Table 1 Characteristics of elderly and non-elderly Patients with HCV

		Non-elderly n = 3661[b] number (%)	Elderly n = 364[b]	P-value[a]
Male		3493 (97%)	352 (98%)	0.23
Race/ethnicity	Non-Hispanic white	2088 (57%)	210 (59%)	
	Non-Hispanic black	1037 (29%)	102 (29%)	
	Hispanic/Latino	359 (10%)	31 (9%)	
	Other	152 (4%)	15 (4%)	0.9
Education HS or less		1632 (45%)	213 (59%)	<0.01
Income < $10,000		1428 (39%)	142 (40%)	0.91
History of injection drug use		2303 (63%)	92 (25%)	<0.01
Recent substance abuse		759 (21%)	32 (9%)	<0.01
Daily alcohol use	None	559 (16%)	100 (30%)	
	<3 drinks/day	232 (7%)	46 (14%)	
	3–6 drinks/day	484 (14%)	58 (17%)	
	>6 drinks/day	2126 (63%)	133 (40%)	<0.01
HIV antibody positive		252 (7%)	10 (3%)	<0.01
Psychiatric disease		696 (20%)	26 (7%)	<0.01
Medical co-morbidities		638 (18%)	137 (38%)	<0.01
Cardiac disease		120 (3%)	52 (15%)	<0.01
Inability to remain compliant		157 (4%)	14 (4%)	0.19
Abnormal ALT		2644 (74%)	226 (63%)	<0.01
Platelets <85 K		193 (5.4%)	25 (7.0%)	0.21
Received liver biopsy		1080 (30.5%)	69 (19.4%)	<0.01
>Stage 2 fibrosis on biopsy		620 (67.4%)	39 (68.4%)	0.87

[a] Chi-square test *P*-value

[b] *N* in each strata may be slightly lower due to missing data

Table 2 Treatment candidacy, preference and initiation rates by elderly status

	Non-elderly n = 3,661[b]	Elderly n = 364[b]	P-value[a]
Treatment candidate per clinician	1,470 (42%)	85 (25%)	<0.01
Treatment candidate per guidelines	934 (26%)	58 (16%)	<0.01
Patient acceptance of treatment[c]	1,097 (77%)	52 (63%)	<0.01
Treatment initiation	719 (20%)	35 (10%)	<0.01

[a] P-value for Chi-square test

[b] N in each strata may be slightly lower due to missing data

[c] Reflects percentage of patients who agreed to be treated out of total number of patients who were deemed treatment candidate by the evaluating physician

Table 3 Unadjusted and adjusted relative odds of being considered a treatment candidate and treatment success associated with being elderly status

	Unadjusted OR (95% CI)	Adjusted OR (95% CI)
Treatment candidacy (by clinician)	0.44 (0.34–0.57)	0.43 (0.30–0.61)[a]
Treatment success (SVR)	1.12 (0.48–2.62)	1.54 (0.46–5.16)[b]

[a] Adjusted for sex, race, education, income, HIV status, abnormal ALT, substance abuse, alcohol use psychiatric co-morbidities, medical co-morbidities, cardiac disease, and non-compliance

[b] Adjusted for sex, race, HCV genotype, viral load, BMI, stage fibrosis

Table 4 Treatment outcomes by elderly status

	Non-elderly n = 719	Elderly n = 35	P-value[a]
Experienced side effects	126 (18%)	9 (25%)	0.22
Early discontinuation	219 (31%)	11 (31%)	0.9
End of treatment viral response	206 (29%)	12 (34%)	0.47
Sustained viral response	131 (18%)	7 (20%)	0.79

[a] P-value for Chi-square test

of patients experienced side-effects and approximately a third discontinued treatment early. Using univariate analysis of the patients treated, there was no difference in the rate of EOTR and SVR between elderly and non-elderly patients (34% vs. 29% and 20% vs. 18%, respectively). This was also confirmed by logistic regression, which showed that elderly patients did not appear to be less likely to achieve SVR compared to non-elderly patients (aOR 1.54; 95% CI: 0.46–5.15), even after adjustment for sex, race, HCV genotype, viral load, BMI, and stage of fibrosis (Table 3).

Discussion

In this national prospective study of patients referred for treatment of chronic hepatitis C at VA Medical Centers, we found that nearly 10% of referred patients were elderly (age > 60 years). The elderly patients who were referred did not appear to have indices to suggest more severe liver disease (platelets, ALT levels, fibrosis on liver biopsy). However, we found that elderly patients had different exclusionary criteria for treatment compared to younger patients: older patients were less likely to have the typical barriers such as substance abuse and psychiatric disorders; however, they were more likely to have medical co-morbidities, such as heart disease. Still, adjusting for these factors and others, we found that older age was independently associated with a lower likelihood of being considered a treatment candidate. In contrast, it appeared that being elderly was not associated with an increased likelihood of experiencing treatment side-effects or experiencing treatment failure.

This is the first study of which we are aware of to look at specific factors associated with treatment candidacy and clinician impressions of treatment candidacy among elderly patients with chronic hepatitis C. Although treatment guidelines do not specifically advise clinicians to withhold treatment for older adults, it is perhaps not surprising that older age was independently associated with a lower likelihood of being considered a treatment candidate and patient acceptance of treatment. Because chronic HCV

is believed to induce liver fibrosis at a slow rate, a life expectancy of at least one or two decades should be a prerequisite for treatment in order for the benefits to be realized. However, since many patients in their 60s and 70s who are in good health can reasonably expect to live 20 years or more, treatment for chronic HCV may still be beneficial for certain patients. It is worth noting that even after adjusting for medical co-morbidities (which should help predict life-expectancy), age was still a predictor of not being considered a treatment candidate in this study.

Our finding that older age was not associated with a lower likelihood of treatment success is similar to what has been observed in prior research. Prior small observational studies conducted in Japan and the United States suggested that treatment with interferon may be equally effective in older compared to younger adults [12–14]. The largest study to date from France observed that SVR was achieved in 45% of the 170 patients 65 years of age or older who were treated with pegylated interferon and ribavirin for their chronic HCV [15]. Despite reports of more frequent cytopenias [16] and early discontinuation [13] reported in the literature, older patients in this study did not appear more likely to suffer side effects or need to discontinue treatment early due to side effects.

This study has important implication for clinical care. In the U.S., the majority of individuals infected with HCV will be entering their 60s within the next few decades [1]. Because patients with chronic hepatitis C are largely asymptomatic, many patients may not become aware of their condition until they interface with the medical system as an older adult. Therefore, the evaluation of elderly patients with HCV for treatment may become a more common scenario in the future. This study demonstrates that clinicians are less likely to treat older patients for HCV. However, it is unclear whether this disparity reflects clinicians' accurate assessments of the life-expectancies of their patients or if physicians may unconsciously have an age bias. Furthermore, data on long-term outcomes associated with treatment are needed for all patients, including older patients. Clinicians who treat patients with HCV may benefit from a clinical framework for evaluating the need for HCV treatment in older adults, similar to what has been developed for cancer screening in older adults [17].

The strengths of this study include the large overall sample size, the inclusion of patients from multiple different medical centers throughout the United States, and its prospective design for data collection. Although an important finding of the paper, the relatively small percentage of elderly patients who ultimately received treatment for HCV (and the use of non-pegylated interferons) resulted in low rates of treatment success, and thus provided limited power for our treatment outcomes analyses. An additional limitation is the fact that this study enrolled patients who had already been referred to specialty clinics for evaluation for treatment. Therefore, the characteristics of older patients in our sample may not reflect the more general population of elderly patients with HCV who may have even higher rates of co-morbidities. However, since all patients came from this referral population, comparisons between older and younger patients should remain valid. Older patients who were treated may have been more aggressively screened for other health issues, which may have resulted in better treatment outcomes in that group. Finally, because this study was conducted among recipients of healthcare within the VA healthcare system, results may lack generalizeability.

In summary, this study of users of the VA healthcare found that only 10% of patients over the age of 60 years who were referred for management of HCV were treated. Older age was independently associated with not being considered a treatment candidate, even after adjustment for co-morbidities. Although only a small number of elderly patients were treated, results suggest that side effects, early discontinuation and, most importantly, treatment outcomes are similar between elderly and non-elderly patients. Further research is needed to determine how clinicians are using age to factor into their decisions on treatment candidacy.

Acknowledgements *Other Members of the VA-HCV-001 Study Group*: Richard Moseley, Ann Arbor; Curt Hagedorn, Atlanta; David Johnson, Bay Pines; Marcos Pedrosa, Boston; Ayse Aytaman, Brooklyn; Charles Mendenhall, Cincinnati; Bhupinder Anand, Houston; Paul King, Columbia (MO); Franz Simon, Denver; Warren Schmidt, Iowa City; Ke-Qin Hu, Loma Linda; Timothy Morgan, Long Beach; Bradford Waters, Memphis; Lennox Jeffers, Miami; Samuel Ho, Minneapolis; Joseph Awad, Nashville; Ramsey Cheung, Palo Alto; Kyong-Mi Chang, Philadelphia; Jawad Ahmad, Pittsburgh; Anna Sasaki, Portland; Doris Strader/Ken Hirsch, Washington DC. This study was funded in part by the VA National HIV/Hepatitis Program and Schering-Plough Corporation. The authors also wish to thank Roche Molecular Systems, Inc. for providing Amplicor kits for HCV RNA testing for this study.

References

1. Armstrong GL, Wasley A, Simard EP et al (2006) The prevalence of hepatitis C virus infection in the United States, 1999 through 2002. Ann Intern Med 144:705–714
2. NIH Consensus Statement on Management of Hepatitis C: (2002) NIH Consens State Sci Statements 19:1–46
3. Strader DB, Wright T, Thomas DL et al (2004) Diagnosis, management, and treatment of hepatitis C. Hepatology 39:1147–1171
4. Marcellin P, Asselah T, Boyer N (2002) Fibrosis and disease progression in hepatitis C. Hepatology 36:S47–S56
5. Poynard T, Ratziu V, Benmanov Y et al (2000) Fibrosis in patients with chronic hepatitis C: detection and significance. Semin Liver Dis 20:47–55
6. Seeff LB (2002) Natural history of chronic hepatitis C. Hepatology 36:S35–S46

7. Regev A, Schiff ER (2001) Liver disease in the elderly. Gastroenterol Clin North Am 30:547–63, x–xi
8. El-Serag HB, Mason AC (1999) Rising incidence of hepatocellular carcinoma in the United States. N Engl J Med 340:745–750
9. Bini EJ, Brau N, Currie S et al (2005) Prospective multicenter study of eligibility for antiviral therapy among 4,084 U.S. veterans with chronic hepatitis C virus infection. Am J Gastroenterol 100:1772–1779
10. Brau N, Bini EJ, Currie S et al (2006) Black patients with chronic hepatitis C have a lower sustained viral response rate than non-Blacks with genotype 1, but the same with genotypes 2/3, and this is not explained by more frequent dose reductions of interferon and ribavirin. J Viral Hepat 13:242–249
11. VA Centers of Excellence in Hepatitis C Research and Education. VA Treatment Recommendations for Patients with Chronic Hepatitis C: version 1.2. 2-25-1999
12. Horiike N, Masumoto T, Nakanishi K et al (1995) Interferon therapy for patients more than 60 years of age with chronic hepatitis C. J Gastroenterol Hepatol 10:246–249
13. Koyama R, Arase Y, Ikeda K et al (2006) Efficacy of interferon therapy in elderly patients with chronic hepatitis C. Intervirology 49:121–126
14. Van Thiel DH, Friedlander L, Caraceni P, Molloy PJ, Kania RJ (1995) Treatment of hepatitis C virus in elderly persons with interferon alpha. J Gerontol A Biol Sci Med Sci 50:M330–M333
15. Thabut D, Le Calvez S, Thibault V et al (2006) Hepatitis C in 6,865 patients 65 year or older: a severe and neglected curable disease? Am J Gastroenterol 101:1260–1267
16. Nudo C, Gupta S, Alpert E et al Elderly patients are at greater risk for cytopenia during anitviral therapy for hepatitis C. CDDW 2004 abstract, accessed at: http://www.cagacg.org/cddw/cddw2004/abs/abs048.htm
17. Walter LC, Covinsky KE (2001) Cancer screening in elderly patients: a framework for individualized decision making. Jama 285:2750–2756

Long-Term Trends in the Incidence of Heart Failure After Myocardial Infarction

Raghava S. Velagaleti, MD; Michael J. Pencina, PhD; Joanne M. Murabito, MD; Thomas J. Wang, MD; Nisha I. Parikh, MD; Ralph B. D'Agostino, PhD; Daniel Levy, MD; William B. Kannel, MD; Ramachandran S. Vasan, MD

Background—Although mortality after myocardial infarction (MI) has declined in the United States in recent decades, there have been few community-based investigations of the long-term trends in the incidence of heart failure after MI, and their results appear to be conflicting.

Methods and Results—We evaluated 676 Framingham Heart Study participants between 45 and 85 years of age (mean age 67 years, 34% women) who developed a first MI between 1970 and 1999. We assessed the incidence rates of heart failure and of death without heart failure in each of 3 decades (1970 to 1979, 1980 to 1989, and 1990 to 1999). We estimated the multivariable-adjusted risk of events in the latter 2 decades, with the period 1970 to 1979 serving as the referent. The 30-day incidence of heart failure after MI rose from 10% in 1970 to 1979 to 23.1% in 1990 to 1999 (*P* for trend 0.003), whereas 30-day mortality after MI declined from 12.2% (1970 to 1979) to 4.1% (1990 to 1999). The 5-year incidence of heart failure after MI rose from 27.6% in 1970 to 1979 to 31.9% in 1990 to 1999 (*P* for trend 0.02), whereas 5-year mortality after MI declined from 41.1% (1970 to 1979) to 17.3% (1990 to 1999). In multivariable analyses, compared with the period 1970 to 1979, we observed higher 30-day (risk ratio 2.05, 95% confidence interval 1.25 to 3.36) and 5-year (risk ratio 1.74, 95% confidence interval 1.07 to 2.84) risks of heart failure in the decade 1990 to 1999. These trends were accompanied by lower 30-day (risk ratio 0.21, 95% confidence interval 0.09 to 0.47) and 5-year (risk ratio 0.31, 95% confidence interval 0.18 to 0.54) mortality rates in 1990 to 1999.

Conclusions—In the present community-based sample, we observed an increase in the incidence of heart failure in recent decades that paralleled the decrease in mortality after MI. **(*Circulation*. 2008;118:2057-2062.)**

Key Words: heart failure ■ myocardial infarction ■ prognosis ■ risk factors ■ epidemiology

Myocardial infarction (MI) is a leading cause of morbidity and mortality in the United States.[1] Major advances in treatment over the last 4 decades have translated into a considerable decline in mortality rates after MI.[2,3] Heart failure (HF) is a common complication of MI,[1] with the estimated incidence varying from 10% to 40%.[4] Post-MI HF is associated with a markedly elevated risk of death,[5] with an estimated median survival of ≈4 years.[6]

Editorial p 2019
Clinical Perspective p 2062

Given the burden posed by HF after MI, it is important to understand the long-term trends in this condition; however, relatively few population-based studies have evaluated the long-term trends in the incidence of post-MI HF. Furthermore, investigators examining the data from 2 major epidemiological studies on such trends reported apparently conflicting results. Investigators from the Rochester Epidemiological Project reported a 28% decline in the incidence of HF after MI between 1979 and 1994 and concluded that a declining trend existed.[7] Whereas an initial report from the Worcester Heart Attack Study noted a modest decline,[8] a more recent report highlighted an upward trend in incidence of HF after MI between 1975 and 2001 in adjusted analyses.[9] An earlier report from the Framingham Heart Study that evaluated trends in incidence of HF after a Q-wave MI (during the time period between 1950 and 1989) demonstrated no long-term change in incidence.[10]

Divergent longitudinal trends in factors that influence HF after MI may have contributed to the inconsistent results in the literature noted above. Thus, improved survival after an

Continuing medical education (CME) credit is available for this article. Go to http://cme.ahajournals.org to take the quiz.
Received April 3, 2008; accepted August 8, 2008.

From the National Heart, Lung, and Blood Institute's Framingham Heart Study (R.S. Velagaleti, M.J.P., J.M.M., T.J.W., N.I.P., R.B.D., D.L., W.B.K., R.S. Vasan), Framingham, Mass; Division of Cardiology (T.J.W.), Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, Mass; Department of Mathematics and Statistics (M.J.P., R.B.D.), Boston University, Boston, Mass; Center for Population Studies, National Heart, Lung, and Blood Institute, Bethesda, Md (D.L.); and General Internal Medicine (J.M.M.), Preventive Medicine and Cardiology Sections (D.L., R.S. Vasan), Boston Medical Center, Boston University School of Medicine, Boston, Mass.

Guest Editor for this article was Veronique L. Roger, MD, MPH.

The online-only Data Supplement is available with this article at http://circ.ahajournals.org/cgi/content/full/CIRCULATIONAHA.108.784215/DC1.

Correspondence to Ramachandran S. Vasan, MD, Framingham Heart Study, 73 Mount Wayte Ave, Framingham, MA 01702-5803. E-mail vasan@bu.edu

Circulation **is available at http://circ.ahajournals.org** **DOI: 10.1161/CIRCULATIONAHA.108.784215**

MI (due to reperfusion therapy) could lead to an increased pool of "high-risk" patients who are more susceptible to HF.[11] Yet, major therapeutic advances (such as use of angiotensin-converting enzyme inhibitors[12] and angiotensin receptor blockers[13] after MI) may have reduced the occurrence of left ventricular dysfunction and stemmed the susceptibility to HF. A potential limitation of the aforementioned investigations is a lack of analyses that explored concurrently the incidence of HF and of mortality after MI without HF as competing events. Such an analytical strategy could elucidate the relative contributions of the divergent trends noted above to the incidence of HF after MI.

We hypothesized that the incidence of HF after MI may have increased in recent times owing to a lower mortality associated with the condition. We tested this hypothesis by evaluating trends in the incidence of HF and death after a first MI in Framingham Heart Study participants over the time period 1970 to 1999.

Methods

Study Sample and Design

The design and characteristics of the original[14] and the offspring[15] cohorts of the Framingham Heart Study have been detailed elsewhere. Briefly, 5209 participants (original cohort) were enrolled in 1948 and have been evaluated approximately every 2 years. The members of the Framingham offspring cohort, comprising 5124 individuals (children of the original cohort and their spouses), were enrolled in 1971 and have been evaluated approximately every 4 years. Participants from both cohorts who attended routine examinations between the years 1970 and 1999 constituted the sampling frame for the present study. All participants provided written informed consent, and the study was approved by the Institutional Review Board of Boston Medical Center.

Participants 45 to 85 years of age who experienced a first MI between the years 1970 and 1999 were eligible for the present investigation (n=715). After the exclusion of participants who had a history of HF before the index MI (n=39), 676 individuals remained eligible for the present investigation. We grouped participants with a first MI according to the decade of onset of the event (ie, 1970 to 1979, 1980 to 1989, or 1990 to 1999). We chose the 3 decades of interest to capture the prethrombolysis, thrombolysis, and percutaneous coronary intervention eras in the management of MI. We could not extend observations to the most recent decade (2000 onward) because participants with an MI in the present decade would not have a follow-up comparable to that for the previous decades. Also, we wanted to minimize ascertainment bias in the diagnosis of MI and HF as a result of the increased use of highly sensitive biomarkers of acute myocardial necrosis (eg, cardiac troponins) or neurohormonal activation (eg, B-type natriuretic peptide), respectively, in the post-2000 time period.

Only the MI event should have occurred in the decade of interest for the participant to be grouped under that decade. Thus, a participant who developed a first MI in 1978 and then developed HF in 1981 would be classified in the decade 1970 to 1979.

Ascertainment of End Points

All Framingham participants are under surveillance for risk of cardiovascular events (including HF) and death, which are identified from data collected at each Framingham examination and from hospitalization records and physician office visits. An end-point review committee, consisting of 3 physicians, reviews all records and adjudicates the occurrence of events. Criteria for these events have been published previously.[16] Briefly, MI was considered to have occurred when participants demonstrated 2 of 3 criteria: new diagnostic Q waves on ECGs, prolonged ischemic chest discomfort, and elevation of circulating cardiac enzymes that suggested myocardial necrosis.

The Framingham criteria for HF[17] were used to adjudicate episodes of HF after MI. Briefly, a diagnosis of HF requires the presence of 2 major or 1 major and 2 minor criteria. The major criteria include a history of paroxysmal nocturnal dyspnea or orthopnea, presence of jugular venous distention, hepatojugular reflux, pulmonary rales, presence of third heart sound, increasing radiographic cardiomegaly, radiographic evidence of acute pulmonary edema, presence of a third heart sound, and evidence of weight loss >4.5 kg during the first 5 days of treatment for suspected HF. The minor criteria include history of a nocturnal cough, dyspnea on ordinary exertion, presence of bilateral ankle edema, hepatomegaly, heart rate >120 bpm, and radiographic evidence of bilateral pleural effusions or pulmonary vascular congestion. Major or minor criteria were attributed to HF only when no alternative explanation could be found for the symptoms or signs (eg, other medical conditions such as cirrhosis, renal failure, or chronic pulmonary disease).

Statistical Analyses

First, participants with a first MI were classified into 3 groups based on the decade of incidence (ie, 1970 to 1979, 1980 to 1989, and 1990 to 1999). Next, we evaluated the incidence of HF or death (free of HF) on follow-up during the 30-day and 5-year time periods after MI. Mortality free of HF was evaluated as opposed to total mortality because mortality in individuals with both MI and HF may be attributable to HF, and we intended to assess mortality as a competing event to the incidence of HF (ie, people who die free of HF "escape" HF). Third, we performed multivariable analyses comparing the incidences of HF and death (separate analyses for each outcome) over the 30-day and 5-year follow-up periods (separate analyses for each follow-up period) after MI occurrence in each of the 3 decades, with the decade 1970 to 1979 serving as the referent group. Fourth, we repeated analyses evaluating the 5-year incidence of HF and death among participants with an MI who survived beyond 30 days (ie, these analyses paralleled earlier analyses but excluded people who died or developed HF within 30 days). Fifth, we performed additional analyses that evaluated incidence of HF or death at time points between 30 days and 5 years (ie, at 6 months, 1 year, and 2 years after MI).

Poisson regression was used for modeling events during the 30-day period after MI, whereas Cox proportional hazards regression was used for incidence of events over the 5-year period after MI. All regression models adjusted for the following covariates (obtained from the Framingham Heart Study examination that preceded the incident MI event): age, sex, body mass index, smoking status, systolic blood pressure, hypertension treatment, diabetes mellitus, and total cholesterol. Given the changes in ascertainment of MI across the decades,[18] we conducted additional analysis adjusting for the proportion of MIs diagnosed without diagnostic ECG changes (ie, on the basis of cardiac biomarkers and clinical history without diagnostic Q waves) in each decade. All statistical analyses were performed with SAS software version 8.0 (SAS Institute, Cary, NC), and $P<0.05$ was used to denote statistical significance.

The authors had full access to and take full responsibility for the integrity of the data. All authors have read and agree to the manuscript as written.

Results

The study sample consisted of 676 participants with a first occurrence of MI over the 3 decades of observation. The baseline characteristics of study participants are shown in Table 1 according to the decade of MI incidence (data for covariates being obtained from the examination that preceded the onset of MI). Participants with a first MI in the most recent decade were older and were more likely to be women, to have diabetes, and to be taking antihypertensive agents, but they had lower serum cholesterol concentrations and smoking

Table 1. Baseline Characteristics by Decade of Incidence of MI

Clinical Characteristic	Decade of MI Incidence		
	1970–1979 (n=230)	1980–1989 (n=251)	1990–1999 (n=195)
Age, y, mean±SD	63.6±10.1	66.9±10.2	69.9±10.9
Women, %	30.9	31.5	38.5
Body mass index, kg/m^2, mean±SD	27.2±4.43	27.0±4.31	28.1±4.66
Systolic blood pressure, mm Hg, mean±SD	145.2±22.2	139.5±20.5	142.7±21.4
Diastolic blood pressure, mm Hg, mean±SD	83.8±11.8	80.8±10.4	77.9±11.7
Hypertension treatment, %	29.3	36.5	44.6
Total cholesterol, mg/dL, mean±SD	239.7±42.4	235.3±44.3	216.5±41.7
Diabetes mellitus, %	11.5	17.5	21.5
Smoking, %	48.2	38.5	29.5
Q-wave MI, %	75.2	71.7	47.7

rates. A larger proportion of MIs in the 1990s were non–Q-wave MIs (Table 1).

Trends in the Incidence of HF and Death

HF after MI occurred in 165 participants with MI (24.4%), whereas 139 participants (20.6%) died without HF over the 3 decades of observation. Table 2 displays the age- and sex-adjusted rates of incident events (HF and death free of HF) in each decade separately for the initial 30 days after MI and for the 5-year period after the MI.

The Figure displays the age- and sex-adjusted survival free of HF during the initial period after an MI in the 3 decades. The incidence of HF during the 30 days after MI was higher in the decades 1980 to 1989 and 1990 to 1999 than in 1979 to 1979 (P for trend=0.003). In contrast, the incidence of death without HF during the 30 days after MI was lower in these decades (1980 to 1989 and 1990 to 1999) than in the period 1970 to 1979 (P for trend $<$0.0001). The multivariable-adjusted hazards ratio for HF after MI during the 30-day period was ≈2-fold higher in the 1990s than for participants who developed MI in the 1970s. In contrast, the adjusted hazard of death without HF was 50% lower in the 1980s and 80% lower in the 1990s than in the referent decade of the 1970s. When the composite outcome of death or HF after MI was modeled, no statistically significant differences in incidence were found across the 3 decades (Table 2).

For the 5-year post-MI time period (including the 30-day post-MI period), the incidence of HF after MI rose in the 1990s compared with the 1970s (P for trend=0.02), whereas the incidence of death without HF declined over this time period (P for trend $<$0.0001). In multivariable models, the risk of new-onset HF was ≈75% higher in the 1990s than in the 1970s. In comparison, the adjusted risk of death free of HF declined by 50% (1980s) to 70% (1990s) over this time period (Table 2). The trend in the incidence of the composite outcome of death or HF after MI was not statistically significant.

For participants who survived beyond 30 days after MI without HF, no statistically significant difference was found in the incidence of HF over a 5-year period, although the incidence of death without HF was lower in the 1980s and 1990s than in the 1970s (P for trend=0.01). The incidence of the composite outcome of death or HF after MI declined by ≈36%, a finding that was of borderline statistical significance. The test for a statistical interaction between the decade of MI incidence and the timing of post-MI HF (dichotomized at 30 days) was highly significant ($P<0.001$), which confirmed that the effect of decade of MI incidence on the incidence of HF after MI diminished beyond the 30-day period after MI. In additional analyses that adjusted for the proportion of MIs ascertained without diagnostic ECG changes in each decade, the observed trends in post-MI HF incidence across decades remained robust.

The Appendix Table (online-only Data Supplement) presents data on incidence of HF and death at 6 months, 1 year, and 2 years after MI. These data suggest that the higher incidence of HF observed at 30 days in recent decades was maintained in analyses of HF and death after MI at these time points as well.

Discussion

Principal Findings

Our principal findings are 2-fold. First, participants with a first MI had a decreasing trend for mortality free of HF between 1970 and 1999 and a concomitant increasing trend for the incidence of HF. These trends were evident for both the 30-day post-MI period and the 5-year post-MI period. Additional analyses suggested that these trends were not influenced by the increasing trend for ascertainment of MI based on biomarkers (and the resultant potential change in case mix of MI). Also, we consistently used the same set of criteria for the ascertainment of HF across the decades; however, trends in the ascertainment of HF based on a greater performance of imaging tests or a greater diagnostic suspicion in more recent decades may have contributed to the present finding of a greater incidence of HF in the 1990s.[19] Second, in the participants who survived beyond the 30-day post-MI period without HF, we did not observe any temporal trends in the incidence of HF after MI. These data suggest that the rising trend in incidence of HF after MI was driven largely by the trend for increased incidence noted for the 30-day post-MI period in recent decades. It is conceivable that our observations are consistent with the well-acknowledged lower mortality and better myocardial salvage of individuals with an MI in the 1990s.[20] Survivors of MI have residual myocardial damage and a higher risk of developing HF.[21] Improved survival in recent decades may have contributed to an increase in the pool of people at risk for developing HF, thereby explaining the increasing trend in the incidence of HF after MI from 1970 to 1999; such individuals may have experienced higher mortality rates in the earlier decades.

Table 2. Age- and Sex-Adjusted Event Rates After MI and Adjusted Relative Risk of Events Across Decades

	1970–1979	1980–1989	1990–1999	*P* for Trend
30-Day events				
CHF				0.003
No. of events/No. at risk (%)	23/230 (10.0)	36/251 (14.3)	45/195 (23.1)	
Event rate (95% CI)*	11.8 (7.5–18.6)	14.6 (10.1–21.2)	19.2 (13.6–27.0)	
Risk ratio (95% CI)	Referent	1.33 (0.80–2.22)	2.05 (1.25–3.36)	
Death without CHF				<0.0001
No. of events/No. at risk (%)	28/230 (12.2)	19/251 (7.6)	8/195 (4.1)	
Event rate (95% CI)*	15.0 (9.2–24.5)	7.6 (4.4–13.2)	3.4 (1.6–7.1)	
Risk ratio (95% CI)	Referent	0.51 (0.29–0.90)	0.21 (0.09–0.47)	
CHF or death				0.79
No. of events/No. at risk (%)	51/230 (22.2)	55/251 (21.9)	53/195 (27.2)	
Event rate (95% CI)*	26.6 (19.8–35.7)	22.2 (16.8–29.5)	22.4 (16.9–29.8)	
Risk ratio (95% CI)	Referent	0.85 (0.61–1.19)	0.95 (0.67–1.34)	
5-Year events (including 30-day events)				
CHF				0.02
No. of events/No. at risk (%)	45/230 (19.6)	54/251 (21.5)	66/195 (33.9)	
Event rate (95% CI)*	27.6 (18.8–35.0)	24.6 (17.3–30.9)	31.9 (23.5–39.0)	
Hazards ratio (95% CI)	Referent	1.05 (0.68–1.63)	1.74 (1.07–2.84)	
Death without CHF				<0.0001
No. of events/No. at risk (%)	66/230 (28.7)	47/251 (18.7)	26/195 (13.3)	
Event rate (95% CI)*	41.1 (30.8–48.9)	23.9 (16.2–30.4)	17.3 (10.1–23.7)	
Hazards ratio (95% CI)	Referent	0.48 (0.32–0.73)	0.31 (0.18–0.54)	
CHF or death				0.25
No. of events/No. at risk (%)	111/230 (48.3)	101/251 (40.2)	92/195 (47.2)	
Event rate (95% CI)*	54.0 (45.6–60.3)	40.7 (33.3–46.9)	41.6 (33.8–48.2)	
Hazards ratio (95% CI)	Referent	0.72 (0.53–0.97)	0.83 (0.58–1.18)	
5-Year events in people surviving 30 days				
CHF				0.99
No. of events/No. at risk (%)	22/179 (12.3)	18/196 (9.2)	21/142 (14.8)	
Event rate (95% CI)*	17.0 (8.1–24.6)	10.6 (4.6–16.0)	14.6 (6.9–21.5)	
Hazards ratio, 95% CI	Referent	0.69 (0.34–1.39)	1.02 (0.47–2.21)	
Death without CHF				0.01
No. of events/No. at risk (%)	38/179 (21.2)	28/196 (14.3)	18/142 (12.7)	
Event rate (95% CI)*	28.3 (17.7–36.6)	14.9 (8.4–20.7)	12.2 (5.8–17.9)	
Hazards ratio (95% CI)	Referent	0.53 (0.31–0.90)	0.43 (0.21–0.87)	
CHF or death				0.06
No. of events/No. at risk (%)	60/179 (33.5)	46/196 (23.5)	39/142 (27.5)	
Event rate (95% CI)*	38.6 (28.5–46.4)	23.3 (15.8–29.7)	24.2 (16.0–31.2)	
Hazards ratio (95% CI)	Referent	0.59 (0.39–0.91)	0.64 (0.38–1.07)	

Multivariable models adjust for the following covariates: age, sex, body mass index, systolic blood pressure, hypertension treatment, total cholesterol, diabetes mellitus, and smoking.

*All event rates are age and sex adjusted per 100 people.

Of note, a previous Framingham report that evaluated trends in the incidence of HF (of any cause) in the time period 1950 to 1999 demonstrated that incidence is stable in men but may be decreasing in women.[22] In the present report, we demonstrate an increasing trend in post-MI HF over the time period from 1970 to 1999. It may be important to analyze trends in incidence of HF due to specific causes, because temporal patterns may vary on the basis of the cause of HF. For instance, it is conceivable that HF caused by hypertension may have declined in more recent decades owing to better control of high blood pressure, thereby explaining the decline in overall incidence of HF in women (in whom the contribution of high blood pressure is greater relative to that of MI[23]).

Trends in HF After MI Incidence: Comparison With Previous Literature

The results of the present investigation vary from those of some other reports in the published literature. Differences in

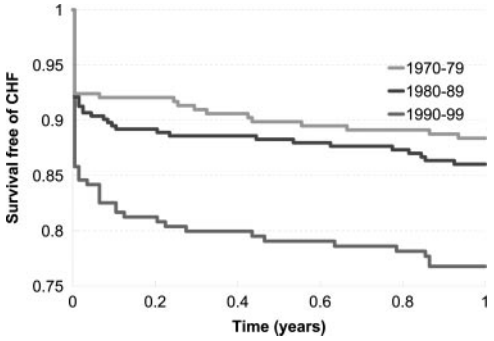


Figure. Age- and sex-adjusted survival free of HF after MI in the 3 decades of interest during the year after the incident MI. CHF indicates chronic HF.

the case mix in the study samples (inclusion of first versus recurrent MI; incident versus prevalent HF), varying durations of follow-up (in-hospital versus short- and long-term follow-up after MI), and distinctions in the time periods of observation (inclusion of the early and late 1990s versus analysis of data from the early 1990s) may have contributed to the apparently dissimilar findings across these studies, as detailed below.

An investigation of the Olmsted County, Minnesota, population by Hellerman et al[7] that used a study sample and design similar to the present study (prospective cohort with incident MIs, no prevalent HF, and mean follow-up of 7.6 years) reported a 28% decline in the incidence of post-MI HF between 1979 and 1994. A major difference between that report and the present report is the time period under study (1979 to 1994 versus 1970 to 1999). As noted by Goldberg et al,[9] the use of primary percutaneous intervention as a treatment for MI became more common in the late 1990s. It is possible that the increased survival of "sicker" patients with MI because of the efficacy of primary percutaneous intervention led to an accrual of more susceptible people in the latter half of the 1990s, which could explain why we observed an increasing trend in post-MI HF by studying people up to 1999.

Several reports from the Worcester Heart Attack Study (WHAS), a longitudinal, community-based surveillance study, also have evaluated temporal trends in the incidence of HF after MI. These reports were based on abstraction of hospitalization records of patients with MI in the Worcester Standard Metropolitan Area in different time periods, focused on incidence of HF during the initial hospital stay, evaluated all patients hospitalized with MIs (both first and recurrent), and included patients with or without prevalent HF. In contrast, the present investigation evaluated both short- and long-term incidence (both during and beyond the initial hospital stay) of HF after a first MI. An initial analysis of 20 years of data (up to the year 1995) by Spencer et al[8] from WHAS showed an inconsistent trend for HF incidence. The 25-year analysis (up to year 2001) reported by Goldberg et al[9] showed unadjusted HF incidence rates that were similar across the time periods evaluated; however, after statistical adjustment for age, sex, prevalent coronary artery disease, and MI order and type, the investigators observed an increasing trend in HF incidence in recent decades. As noted above, it is possible that the latter findings are consistent with the present observations because of the similarity in the time periods studied.

Strengths and Limitations

The present study extends prior observations in several respects. We evaluated both in-hospital HF and events after the index MI over a period of 5 years. Furthermore, we concomitantly evaluated trends in both incidence of HF and death (free of HF) to assess how changing case-fatality rates may influence the incidence of post-MI HF.

Nevertheless, several limitations of the present investigation must be noted. We did not model temporal trends in the use of specific treatments (such as use of angiotensin-converting enzyme inhibitors, angiotensin receptor blockers, aldosterone antagonists, or β-blockers) or revascularization procedures for participants with an MI. It is more likely that higher proportions of participants in the latter decades received these therapies than participants in the decade of 1970 to 1979[20]; however, treatments that improve post-MI survival have also been shown to decrease the incidence of HF. Another limitation of the present study is the lack of information on the type of HF (systolic versus diastolic) or the occurrence of left ventricular systolic dysfunction after MI. Lastly, the present study sample is predominantly white and of European ancestry, and caution must be exercised in generalizing these results to other ethnicities.

Conclusions

Our longitudinal observations on a large, community-based sample demonstrate reciprocal trends of a decrease in mortality after MI accompanied by an increase in incidence of HF in more recent years (1990s) relative to the 1970s. Greater salvage of high-risk MI patients in recent time periods may contribute to these trends.

Sources of Funding

This work was supported by National Institutes of Health/National Heart, Lung, and Blood Institute contract No. N01-HC-25195 and 2K24HL4334 (Dr Vasan).

Disclosures

None.

References

1. Rosamond W, Flegal K, Friday G, Furie K, Go A, Greenlund K, Haase N, Ho M, Howard V, Kissela B, Kittner S, Lloyd-Jones D, McDermott M, Meigs J, Moy C, Nichol G, O'Donnell CJ, Roger V, Rumsfeld J, Sorlie P, Steinberger J, Thom T, Wasserthiel-Smoller S, Hong Y. Heart disease and stroke statistics: 2007 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation*. 2007;115:e69–e171.
2. Rosamond WD, Chambless LE, Folsom AR, Cooper LS, Conwill DE, Clegg L, Wang CH, Heiss G. Trends in the incidence of myocardial infarction and in mortality due to coronary heart disease, 1987 to 1994. *N Engl J Med*. 1998;339:861–867.
3. Fox CS, Evans JC, Larson MG, Kannel WB, Levy D. Temporal trends in coronary heart disease mortality and sudden cardiac death from 1950 to 1999: the Framingham Heart Study. *Circulation*. 2004;110:522–527.
4. Weir RA, McMurray JJ, Velazquez EJ. Epidemiology of heart failure and left ventricular systolic dysfunction after acute myocardial infarction: prevalence, clinical characteristics, and prognostic importance. *Am J Cardiol*. 2006;97:13F–25F.
5. Lewis EF, Moye LA, Rouleau JL, Sacks FM, Arnold JM, Warnica JW, Flaker GC, Braunwald E, Pfeffer MA. Predictors of late development of

heart failure in stable survivors of myocardial infarction: the CARE study. *J Am Coll Cardiol*. 2003;42:1446–1453.
6. Hellermann JP, Jacobsen SJ, Redfield MM, Reeder GS, Weston SA, Roger VL. Heart failure after myocardial infarction: clinical presentation and survival. *Eur J Heart Fail*. 2005;7:119–125.
7. Hellermann JP, Goraya TY, Jacobsen SJ, Weston SA, Reeder GS, Gersh BJ, Redfield MM, Rodeheffer RJ, Yawn BP, Roger VL. Incidence of heart failure after myocardial infarction: is it changing over time? *Am J Epidemiol*. 2003;157:1101–1107.
8. Spencer FA, Meyer TE, Goldberg RJ, Yarzebski J, Hatton M, Lessard D, Gore JM. Twenty year trends (1975–1995) in the incidence, in-hospital and long-term death rates associated with heart failure complicating acute myocardial infarction: a community-wide perspective. *J Am Coll Cardiol*. 1999;34:1378–1387.
9. Goldberg RJ, Spencer FA, Yarzebski J, Lessard D, Gore JM, Alpert JS, Dalen JE. A 25-year perspective into the changing landscape of patients hospitalized with acute myocardial infarction (the Worcester Heart Attack Study). *Am J Cardiol*. 2004;94:1373–1378.
10. Guidry UC, Evans JC, Larson MG, Wilson PW, Murabito JM, Levy D. Temporal trends in event rates after Q-wave myocardial infarction: the Framingham Heart Study. *Circulation*. 1999;100:2054–2059.
11. Gheorghiade M, Bonow RO. Chronic heart failure in the United States: a manifestation of coronary artery disease. *Circulation*. 1998;97:282–289.
12. Kober L, Torp-Pedersen C, Carlsen JE, Bagger H, Eliasen P, Lyngborg K, Videbaek J, Cole DS, Auclert L, Pauly NC; Trandolapril Cardiac Evaluation (TRACE) Study Group. A clinical trial of the angiotensin-converting-enzyme inhibitor trandolapril in patients with left ventricular dysfunction after myocardial infarction. *N Engl J Med*. 1995;333:1670–1676.
13. Pfeffer MA, McMurray JJ, Velazquez EJ, Rouleau JL, Kober L, Maggioni AP, Solomon SD, Swedberg K, Van de Werf F, White H, Leimberger JD, Henis M, Edwards S, Zelenkofske S, Sellers MA, Califf RM; Valsartan in Acute Myocardial Infarction Trial Investigators. Valsartan, captopril, or both in myocardial infarction complicated by heart failure, left ventricular dysfunction, or both [published correction appears in *N Engl J Med*. 2004;350:203]. *N Engl J Med*. 2003;349:1893–1906.
14. Dawber TR, Meadors GF, Moore FE Jr. Epidemiological approaches to heart disease: the Framingham Study. *Am J Public Health Nations Health*. 1951;41:279–281.
15. 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.
16. Kannel WB, Wolf PA, Garrison RJ, eds. *Some Risk Factors Related to the Annual Incidence of Cardiovascular Disease and Death in Pooled Repeated Biennial Measurements. Section 34: The Framingham Heart Study: 30 Year Follow-Up*. Bethesda, Md: National Institutes of Health; 1987. NIH publication No. 87-2703. 1987.
17. McKee PA, Castelli WP, McNamara PM, Kannel WB. The natural history of congestive heart failure: the Framingham study. *N Engl J Med*. 1971;285:1441–1446.
18. Roger VL. Epidemiology of myocardial infarction. *Med Clin North Am*. 2007;91:537–552.
19. Lopez-Jimenez F, Goraya TY, Hellermann JP, Jacobsen SJ, Reeder GS, Weston SA, Roger VL. Measurement of ejection fraction after myocardial infarction in the population. *Chest*. 2004;125:397–403.
20. Rogers WJ, Canto JG, Lambrew CT, Tiefenbrunn AJ, Kinkaid B, Shoultz DA, Frederick PD, Every N. Temporal trends in the treatment of over 1.5 million patients with myocardial infarction in the US from 1990 through 1999: the National Registry of Myocardial Infarction 1, 2 and 3. *J Am Coll Cardiol*. 2000;36:2056–2063.
21. Sutton MG, Sharpe N. Left ventricular remodeling after myocardial infarction: pathophysiology and therapy. *Circulation*. 2000;101:2981–2988.
22. Levy D, Kenchaiah S, Larson MG, Benjamin EJ, Kupka MJ, Ho KK, Murabito JM, Vasan RS. Long-term trends in the incidence of and survival with heart failure. *N Engl J Med*. 2002;347:1397–1402.
23. Levy D, Larson MG, Vasan RS, Kannel WB, Ho KK. The progression from hypertension to congestive heart failure. *JAMA*. 1996;275:1557–1562.

CLINICAL PERSPECTIVE

Mortality due to myocardial infarction (MI) has decreased in recent decades; however, few community-based epidemiological investigations have addressed the long-term trends in the incidence of heart failure after MI. We evaluated trends in the incidence of heart failure after MI in the time period 1970 to 1999 in the Framingham Heart Study cohort. We related the decade of MI incidence to the occurrence of heart failure in the early (within 30 days of MI) and late (after 30 days and up to 5 years) post-MI periods and to the incidence of death free of heart failure. We observed a striking increase in the incidence of heart failure after MI in the decade 1990 to 1999 (compared with the decade 1970 to 1979), accompanied by a decrease in the incidence of death without heart failure after MI over the same time period. We conclude that the increase in heart failure incidence after MI in recent decades was explained primarily by increases in the early post-MI period, in part due to a major decrease in mortality during this period in recent decades. The present data are consistent with the notion that a greater salvage of high-risk patients in recent time periods may have contributed to the observed trends in post-MI heart failure.

JOURNAL OF PALLIATIVE MEDICINE
Volume 11, Number 5, 2008

DOI: 10.1089/jpm.2007.0172

Overcoming Educational Barriers for Advance Care Planning in Latinos with Video Images

ANGELO E. VOLANDES, M.D.,[1] MIGUEL ARIZA, M.D.,[2] ELMER D. ABBO, M.D., J.D.,[3] and MICHAEL PAASCHE-ORLOW, M.D., M.A., M.P.H.[2]

ABSTRACT

***Background*: Studies of end-of-life care have shown that Latino patients want more aggressive care compared to white patients. While this has been attributed to aspects of ethnicity, national origin, and religion, it is possible that limited education might obscure the true relationship between Latino patients and their end-of-life care preferences.**

***Methods*: Spanish-speaking subjects presenting to their primary care doctors were asked their preferences for end-of-life care before watching a video of advanced dementia. Subjects then viewed a 2-minute video of a patient with advanced dementia and were asked again about their preferences. Unadjusted and adjusted logistic regression models were fit using stepwise algorithms to examine factors related to preferences.**

***Results*: A total of 104 subjects completed the interview. Before seeing the video, 42 (40%) subjects preferred comfort care; 43 (41%) desired life-prolonging care; 11 (11%) chose limited care; and 8 (8%) were unsure of their preferences. Subject preferences changed significantly after the video: 78 (75%) of the subjects chose comfort care; 8 (8%) desired life-prolonging care; 14 (13%) chose limited care; and, 4 (4%) were unsure of their preferences ($p < 0.001$). Unadjusted and adjusted analyses revealed a statistically significant difference regarding prevideo preferences based on educational level. After the video, differences in preferences based on educational level disappeared.**

***Conclusions*: Educational level was an independent predictor of end-of-life preferences after hearing a verbal description of advanced dementia. After viewing a video of a patient with advanced dementia there were no longer any differences in the distribution of preferences according to educational level. These findings suggest that educational level is an important variable to consider in research and in patient care when communicating about end-of-life care preferences. While attention to patients' culture is important, it is also important to avoid ascribing choices to culture that may actually reflect inadequate comprehension. Attention to communication barriers with techniques like the video used in the current study may help ensure optimal end-of-life care for Latino patients irrespective of educational level.**

INTRODUCTION

LATINO PATIENTS make up the largest minority group in our health care system. There has been relatively little research exploring the end-of-life care preferences for Latino patients compared to white patients. A few studies that included some Latino patients suggest that Latino ethnicity is predictive of end-of-life preferences in favor of life-prolongation.[1–6] Because decision making at the end of life is often

[1]General Medicine Unit, Massachusetts General Hospital, Boston, Massachusetts.
[2]Department of Medicine, Boston Medical Center, Boston, Massachusetts.
[3]Section of General Internal Medicine, Department of Medicine, University of Chicago, Chicago, Illinois.

complex, limited education may obscure patients' preferences. In the end-of-life context, barriers posed by education are surmountable while true value differences should be respected.

Previous studies that suggested Latino ethnicity is predictive of end-of-life preferences favoring life-prolongation did not adequately address the possible confounding role of education. In addition, if it were found that people with limited education do not adequately comprehend future health states to make informed decisions about end-of-life preferences, there are multiple education techniques to overcome barriers posed by limited educational level.

Visual images have been shown to improve communication of information with patients.[7–12] Prior interventions to improve decision making have included print, audio, video and Web-based pictorial materials,[9,13–16] with mixed results.[17] Video images allow patients to envision health states in a manner not easily captured with verbal communication and can both engage people and efficiently communicate information about the experience of illness. Making sure that patients understand disease states at the end of life is a key element to ensuring that optimal care is delivered at the end of life.

In order to explore end-of-life care preferences in Latino patients, we used the model of advanced dementia to study whether care preferences were predicted by educational level. We hypothesized that after hearing a verbal description of advanced dementia, subjects' preferences for care at the end of life would be independently predicted by educational level. Additionally, we hypothesized that a video decision aid of a patient with advanced dementia would overcome communication barriers posed by limited education resulting in a change of preferences for end-of-life care toward comfort care.

METHODS

Subjects

Spanish-speaking patients over the age of 40 scheduled to see a general internist at two study sites were eligible to participate. The study sites consisted of two urban primary care clinics serving the Latino community in the greater Boston area, which consists predominantly of Latinos from the Dominican Republic and Puerto Rico. Subjects were excluded if, in the judgment of the physician interviewer, they lacked medical decision-making capacity at the time of the interview, or if they did not speak Spanish. All interviews were conducted by two Spanish-speaking physicians (A.E.V. and M.A.) between December 1, 2005 and January 31, 2007. Approval for the project was granted by the Institutional Review Boards of the affiliated hospitals.

Design

A structured questionnaire used in a previous study on end-of-life care was translated into Spanish, and then back translated into English.[12] Survey questions were generated from a review of the medical ethics literature, and consultations with medical ethics, palliative care, geriatric, and neurology experts. Early versions of the survey were tested with subjects recruited from primary care clinics.

After obtaining verbal informed consent, interviewers defined advanced dementia in simple language, highlighting functional impairments based on the Functional Assessment Staging (FAST) stage 7c criteria including inability to communicate understandably with others, inability to ambulate without assistance, and inability to feed oneself.[18] Refer to Appendix A for a transcript of the description used. As in our previous studies, we then outlined three levels of medical care and the goals associated with each level: life-prolonging care, limited care, and comfort care.[12] The first level, life-prolonging care, aims to prolong life at any cost. It includes all medically indicated care. The second level, limited care, includes treatments such as hospitalization, antibiotics, and artificial nutrition and hydration, but excludes attempted cardiopulmonary resuscitation and care in the intensive care unit. The third level, comfort care, aims to maximize comfort and to relieve pain. It includes oxygen and analgesics but excludes intravenous therapies and hospitalization unless necessary to provide comfort. Subjects were asked which level they preferred in the event they developed advanced dementia. Subjects not able to select a level of care were documented as "don't know."

Each subject next viewed a 2-minute video of a patient with advanced dementia, which was used in our previous studies. The video depicts the salient features of advanced dementia. The Spanish narrative that accompanies the video is identical to the verbal description used to assess subjects' prevideo preferences. The Spanish film clip used in this study is available online at: ⟨http://homepage.mac.com/avolandes/AlzheimersVideo/iMovieTheater18.html⟩ (an English version of the film is available at www.advancecareplanningvideos.com).

The subject was then asked the same questions as previously regarding preferences for level of medical

care. Answers to sociodemographic questions (age, ethnicity, gender, etc.) were self-reported by the subjects.

Statistical analysis

Bivariate analyses of baseline data were performed to assess the association between various subject characteristics and preferences. Subject characteristics were compared across groups using the chi-square test for categorical variables and the nonparametric Kruskal-Wallis test for continuous variables. Subjects' choices for advanced care planning included: life-prolonging care, limited care, comfort care and don't know. These four responses were divided into comfort care versus aggressive care, which included life-prolonging care, limited care, and do not know. Separate logistic regression models were constructed to examine the association between initial preferences after a verbal description (aggressive care, yes versus no) and factors obtained in our baseline survey, which included demographic items as well as additional attributes obtained in the baseline survey (age, gender, education, marital status, religion, religious attendance, health status, and national origin). Multivariate logistic regression analyses were used to identify factors independently associated with preferences. Stepwise algorithms were used to assess the independent role of each variable. Factors significant at $p < 0.10$ were retained. Factors thus selected were simultaneously entered into the logistic regression model. Another backward stepwise algorithm, using $p < 0.05$, was used to determine the final model. All analyses were carried out using Stata software, version 8 (Stata Corporation, College Station, TX).

RESULTS

Subjects

A total of 104 subjects agreed to be interviewed. Table 1 describes the characteristics of the survey sample. Of the 104 subjects, 40 (38%) were Dominican and 42 (40%) were Puerto Rican. Most of the subjects were Christian (95%), consisting of both Catholic (63%) and other Christian denominations (33%). Many subjects (34%) had only an elementary education, while most (54%) had a high school education; few (9%) had a college education.

Table 1. Characteristcs of Study Subjects

Characteristics	*Total*	*Dominicans*	*Puerto Rican*	*Others*[a]
Number of subjects, no. (%)	104	40 (38)	42 (40)	22 (21)
Age, mean (SD)	**55 (10)**	**53 (12)**	**58 (12)**	**51 (8)**
Women, no. (%)	93 (64)	56 (68)	37 (57)	0.22
Education, no. (%)				
Elementary school	**38 (37)**	**20 (50)**	**16 (38)**	**2 (9)**
High school	**56 (54)**	**16 (40)**	**22 (52)**	**18 (82)**
College or beyond	**10 (9)**	**4 (10)**	**4 (10)**	**2 (9)**
Marital status, no. (%)				
Married	49 (47)	16 (40)	24 (57)	9 (41)
Widowed	6 (6)	2 (5)	3 (7)	1 (5)
Single	38 (37)	15 (38)	12 (29)	11 (50)
Divorced	11 (11)	7 (18)	3 (7)	1 (5)
Religion, no. (%)[b]				
Catholic	65 (63)	27 (68)	26 (62)	12 (55)
Christian (non-Catholic)	34 (33)	12 (30)	13 (31)	9 (41)
Non-Christian	5 (5)	1 (3)	3 (7)	1 (5)
Religious attendance, no. (%)				
2 times per month or more	66 (64)	21 (53)	30 (71)	15 (68)
1 time per month or less	20 (19)	9 (23)	6 (14)	5 (23)
Never	17 (17)	9 (23)	6 (14)	2 (9)
Self-reported health status, no. (%)				
Very healthy	25 (24)	7 (18)	11 (26)	7 (32)
Somewhat healthy	52 (50)	19 (48)	24 (57)	9 (41)
Not healthy	27 (26)	14 (35)	7 (17)	6 (27)

[a]Other national origins included: Peru (6); Guatemala (5); El Salvador (4); Honduras (3); Columbia (4); and, Cape Verde (2).
[b]Other Christian denominations reported: Evangelical (10); Pentecostal (9); Protestant (2); Baptist (2); Seventh Day Adventist (2); Episcopalian (1); Jehovah's Witness (1).
Items in bold reflect significant differences across groups with $p < 0.05$.
SD, standard deviation.

OUTCOMES

After hearing a brief verbal description of advanced dementia, 42 (40%) subjects preferred comfort care; 43 (41%) desired life-prolonging care; 11 (11%) chose limited care; and 8 (8%) were unsure of their preferences (Fig. 1). The subjects' preferences changed significantly after the video: 78 (75%) of the subjects chose comfort care; 8 (8%) desired life-prolonging care; 14 (13%) chose limited care; and, 4 (4%) were unsure of their preferences ($p < 0.001$).

Of the 104 subjects, 51 (49%) subjects (95% confidence interval [CI] 37–66) changed their preferences or became uncertain of their preferences after viewing the video (Fig. 1). Of the 42 subjects who chose comfort care before the video, 1 (2%) subject (95% CI 0–5) changed his mind to life-prolonging care.

Of the 62 subjects who did not initially choose comfort care, 37 (60%) subjects (95% CI 29–45) altered their preferences from aggressive levels of care to comfort care.

Of the 43 subjects who chose life-prolonging care before the video, 26 (60%) subjects (95% CI 19–32), chose comfort care after viewing the video. Among the 11 subjects who initially chose limited care, 5 (45%) subjects (95% CI 2–8) chose comfort care after the video. Of the 8 subjects who were uncertain before the video, 6 (75%) subjects (95% CI 3–8) chose comfort care after the video.

Prevideo preferences for care were associated with educational level but not with national origin, religion or any of the other variables (Table 2). Unadjusted analysis revealed that compared to those subjects with a high school or higher education, subjects with less

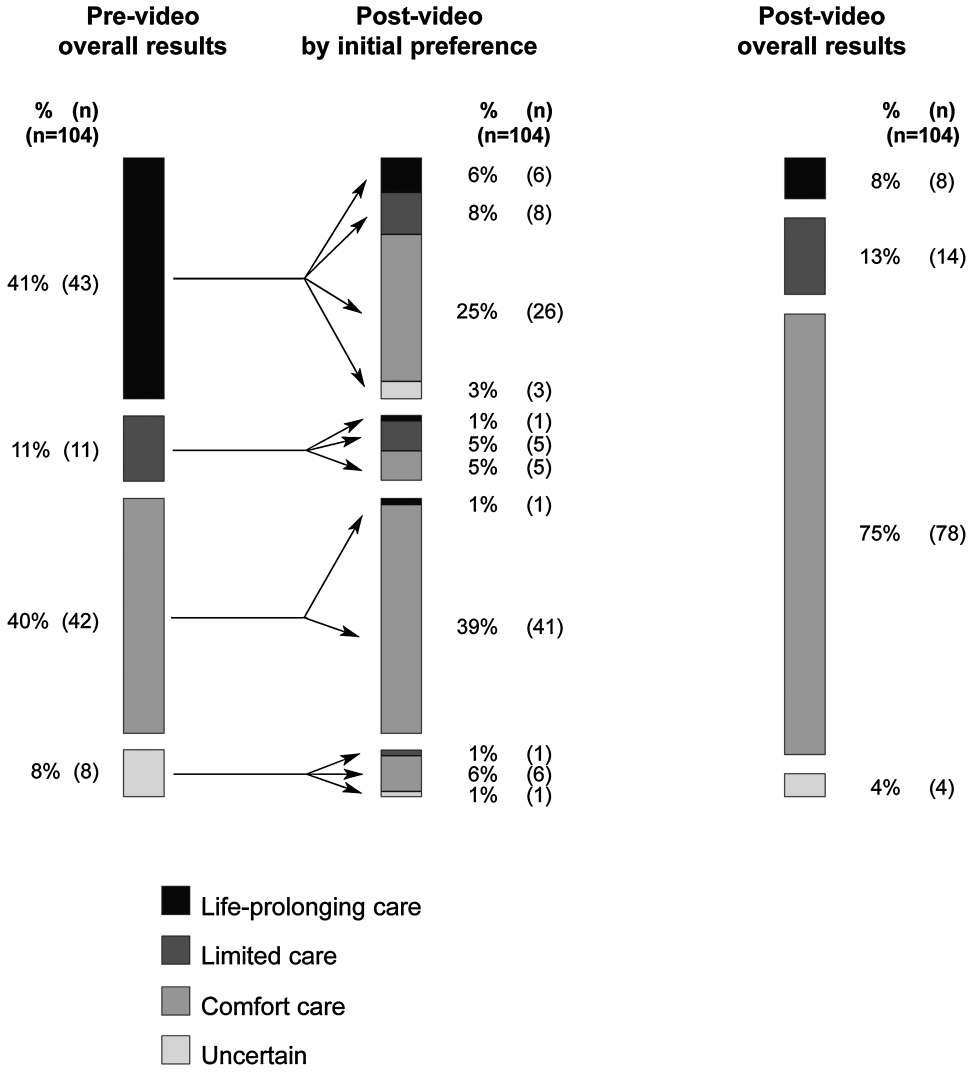


FIG. 1. Patient preferences for level of care before and after the video.

than a high school education were more likely to have preferences for aggressive care (life-prolonging, limited or do not know) after the verbal description, odds ratio (OR) 2.6 (95% CI 1.1–6.3; Table 2). In the multiple logistic regression stepwise model, the effect of education was still significant, OR 2.5 (95% CI 1.1–6.1; Table 3).

After watching the video, all of the significant differences disappeared. The experience of viewing the video was acceptable to the subjects: 73% said they were "comfortable" watching the video; and, 94% thought that using videos for other diseases like cancer would be "helpful."

DISCUSSION

End-of-life preferences for Latino patients were independently predicted by education. Patients with less than a high school education were more likely to have preferences for more aggressive care at the end of life. Furthermore, a video decision aid designed to overcome barriers posed by limited education significantly changed the distribution of preferences at the end of life. After watching a video designed to compensate for limited education, the preferences of Latinos were more likely to be consistent with comfort care.

Several studies have explored facilitating advance care planning options with Latino patients and have concluded that Latinos frequently wish to explore end-of-life care preferences in a family context and have a similar rate of advance care planning to other groups of patients.[19–22] However, few studies have explored what level of treatments Latinos prefer at the end of life, and none of these evaluated the role of education.[1–3,5,6] For example, Blackhall et al.[3] studied 200 elderly Mexican-Americans and found that they were generally more positive about the use of life support and were more likely to personally want such treatments. However, this study did not evaluate the potentially mediating influence of education. Our study suggests that education may mediate the preferences of Latino patients at the end of life and that educational techniques known to compensate for limited education, such as the video used in this project, may surmount barriers posed by limited education. Our findings replicate similar results we have seen in a study of African American patients' end-of-life preferences.[23]

Attention to each patient's values is critical. This study suggests that prior characterizations of Latino preferences for aggressive care may be misleading. If clinicians have the impression that a patient's expressed wishes are concordant with reported commu-

Table 2. Unadjusted Odds Ratios of Initial Preferences for Aggressive Care after Verbal Description

	Initial preferences for aggressive care[a] *after verbal description*	
Characteristics	*Unadjusted OR*	*95% CI*
Ethnicity		
Puerto Rican	0.7	0.3–1.7
Other	1.4	0.5–3.8
Dominican	1.0	
Education		
High school or less	2.6	1.1–6.3
College or beyond	1.0	
Gender		
Male	1.8	0.8–4.2
Female	1.0	
Marital status		
Nonmarried	1.0	
Married	0.64	0.3–1.4
Religion		
Christian (non-Catholic)	1.3	0.6–3.1
Other	2.6	0.4–16.4
Catholic	1.0	
Religious attendance		
≤1 month	0.7	0.2–2.7
≥2 month	0.5	0.2–1.4
Never	1.0	
Health status		
Not healthy	0.5	0.2–1.3
Somewhat	0.5	0.2–1.6
Very healthy	1.0	

[a]Aggressive care includes those subjects who chose one of the following: life-prolonging care, limited care or do not know.
OR, odds ratio; CI, confidence interval.

Table 3. Adjusted Odds Ratios of Initial Preferences for Aggressive Care after Verbal Description

	Initial preferences for aggressive care[b] *after verbal description*	
Characteristics[a]	*Unadjusted OR*	*95% CI*
Education		
High school or less	2.5	1.0–6.1
College or beyond	1.0	
Religious attendance		
≥2 month	0.5	0.2–1.1

[a]Other characteristics (age, gender, religion health status and ethnicity) were excluded from the stepwise model as they were not related to the outcome at a $p < 0.10$.
[b]Aggressive care includes those subjects who chose one of the following: life-prolonging care, limited care or do not know.
OR, odds ratio; CI, confidence interval.

nity norms, they may overlook the roles of good communication and patient education. Indeed, our results indicate that overcoming communication barriers posed by limited educational level should become the focus of clinical care and research at the end of life. Video decision aids are a powerful tool to overcome barriers due to low educational levels. Video is widely accessible and is a medium that can engage people in ways not easily accomplished with words.

Our study has several important limitations. This study asked about preferences in the context of a study in which subjects' decisions were not binding. Additionally, preferences for level of care at the end of life may change over the course of time. Our study looked at the preferences of subjects at a particular point in time. Future work looking at the stability of preferences over time would be helpful, especially in the context of completing advance directives or more formalized advance care planning that could incorporate decision aids such as the video used in the current study.

Video is a powerful medium that can be manipulated to sway patient perspectives. We studied the use of a single video clip of a white patient that was designed to portray advanced dementia in order to overcome barriers posed by low educational level. Future studies using other video clips that varied features such as the ethnicity of the patient in the video would be important.

Our study asked questions about preferences if in a state of advanced dementia, the most common end-of-life scenario. Preferences for end-of-life care in other disease states may be different. Other studies using disease states such as a persistent vegetative state and severe stroke may be fruitful.

Finally, this study primarily looked at Latinos from the Boston area, who are predominantly from the Dominican Republic and Puerto Rico. Our study did not include the largest group of Latinos in this country, Mexicans. Exploring the preferences of Mexicans and other large Latino groups would be of great interest. In addition, our sample was drawn from the metro Boston area. Latinos from other parts of the United States may have different attitudes towards the use of video for decision making at the end of life. These findings need to be reproduced in a nationally representative sample.

Delivering high-quality end-of-life care entails understanding and respecting each patient's values. Our study suggests that inadequate doctor–patient communication may pose a significant barrier to ascertaining patients' values. Video decision aids that overcome barriers posed by low educational level, along with improved verbal communication, may improve the delivery of high-quality end-of-life care to Latino patients. It is the responsibility of the health care system to ensure that the delivery of care at the end of life is consistent with the patient's values and not a result of miscommunication or cultural stereotyping.

REFERENCES

1. Eleazer GP, Hornung CA, Egbert CB, Egbert JR, Eng C, Hedgepeth J, McCann R, Strothers H 3rd, Sapir M, Wei M, Wilson M: The relationship between ethnicity and advance directives in a frail older population. J Am Geriatr Soc 1996;44:938–943.
2. Caralis PV, Davis B, Wright K, Marcial E: The influence of ethnicity and race on attitudes toward advance directives, life-prolonging treatments, and euthanasia. J Clin Ethics. Summer 1993;4(2):155–165.
3. Blackhall LJ, Frank G, Murphy ST, Michel V, Palmer JM, Azen SP: Ethnicity and attitudes towards life sustaining technology. Soc Sci Med 1999;48:1779–1789.
4. Kwak J, Haley WE: Current research findings on end-of-life decision making among racially or ethnically diverse groups. Gerontologist 2005;45:634–641.
5. Duffy SA, Jackson FC, Schim SM, Ronis DL, Fowler KE: Racial/ethnic preferences, sex preferences, and perceived discrimination related to end-of-life care. J Am Geriatr Soc 2006;54:150–157.
6. Balboni TA, Vanderwerker LC, Block SD, Paulk ME, Lathan CS, Peteet JR, Prigerson HG: Religiousness and spiritual support among advanced cancer patients and associations with end-of-life treatment preferences and quality of life. J Clin Oncol 2007;25:555–560.
7. Frosch DL, Kaplan RM, Felitti VJ: A randomized controlled trial comparing Internet and video to facilitate patient education for men considering the prostate specific antigen test. J Gen Intern Med 2003;18:781–787.
8. Morgan MW, Deber RB, Llewellyn-Thomas HA, Gladstone P, Cusimano RJ, O'Rourke K, Tomlinson G, Detsky AS: Randomized, controlled trial of an interactive videodisc decision aid for patients with ischemic heart disease. J Gen Intern Med 2000;15:685–693.
9. Edwards A, Elwyn G, Mulley A: Explaining risks: Turning numerical data into meaningful pictures. BMJ 2002; 324:827–830.
10. Houts PS, Shankar S, Klassen AC, Robinson EB: Use of pictures to facilitate nutrition education for low-income African American women. J Nutr Educ Behav 2006;38: 317–318.
11. Ditto PH, Danks JH, Smucker WD, Bookwala J, Coppola KM, Dresser R, Fagerlin A, Gready RM, Houts RM, Lockhart LK, Zyzanski S: Advance directives as acts of communication: A randomized controlled trial. Arch Intern Med 2001;161:421–430.
12. Volandes AE, Lehmann LS, Cook EF, Shaykevich S, Abbo ED, Gillick MR: Using video images of dementia in advance care planning. Arch Intern Med 2007;167:828–833.

13. Woolf SH, Chan EC, Harris R, Braddock CH 3rd, Kaplan RM, Krist A, O'Connor AM, Tunis S: Promoting informed choice: Transforming health care to dispense knowledge for decision making. Ann Intern Med 2005;143:293–300.
14. Mitchell SL, Tetroe J, O'Connor AM: A decision aid for long-term tube feeding in cognitively impaired older persons. J Am Geriatr Soc 2001;49:313–316.
15. O'Connor AM, Stacey D, Rovner D, Holmes-Rovner M, Tetroe J, Llewellyn-Thomas H, Entwistle V, Rostom A, Fiset V, Barry M, Jones J: Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2001(3):CD001431.
16. Barry MJ: Health decision aids to facilitate shared decision making in office practice. Ann Intern Med 2002;136:127–135.
17. Flory J, Emanuel E: Interventions to improve research participants' understanding in informed consent for research: A systematic review. JAMA 2004;292:1593–1601.
18. Sclan SG, Reisberg B: Functional assessment staging (FAST) in Alzheimer's disease: Reliability, validity, and ordinality. Int Psychogeriatr 1992;4(Suppl 1):55–69.
19. Morrison RS, Meier DE: High rates of advance care planning in New York City's elderly population. Arch Intern Med 2004;164:2421–2426.
20. Froman RD, Owen SV: Randomized study of stability and change in patients' advance directives. Res Nurs Health 2005;28:398–407.
21. Born W, Greiner KA, Sylvia E, Butler J, Ahluwalia JS: Knowledge, attitudes, and beliefs about end-of-life care among inner-city African Americans and Latinos. J Palliat Med 2004;7:247–256.
22. Phipps EJ, True G, Murray GF: Community perspectives on advance care planning: Report from the Community Ethics Program. J Cult Divers 2003;10:118–123.
23. Volandes A, Paasche-Orlow M, Gillick MR, Cook EF, Shaykevich S, Abbo ED, Lehmann L: Health literacy not race predicts end-of-life preferences. J Gen Intern Med 2007;22:5.

Address reprint requests to:
Angelo E. Volandes, M.D.
General Medicine Unit
Massachusetts General Hospital
50 Staniford Street, 9th Floor
Boston, MA 02114

E-mail: avolandes@partners.org

Appendix A. Narrative Describing Advanced Dementia

"I am going to describe to you an illness called advanced dementia, like advanced Alzheimer's dementia, that you may or may not be familiar with. Advanced dementia is an incurable disease of the brain in which one is not able to communicate with others. People in advanced dementia are not able to move around or walk, get out of bed independently, eat by oneself, or communicate understandably with others. People with advanced dementia often have difficulty chewing or swallowing, and require assistance with feeding oneself. Advanced dementia is an incurable disease and most commonly occurs after many years of Alzheimer's disease or as the result of strokes. People are not able to answer any questions or tell you about themselves."

JOURNAL OF PALLIATIVE MEDICINE
Volume 11, Number 5, 2008

DOI: 10.1089/jpm.2007.0224

Health Literacy not Race Predicts End-of-Life Care Preferences

ANGELO E. VOLANDES, M.D.,[1] MICHAEL PAASCHE-ORLOW, M.D., M.A., M.P.H.,[2] MURIEL R. GILLICK, M.D.,[3] E.F. COOK, Sc.D.,[4] SHIMON SHAYKEVICH, M.Sc.,[4] ELMER D. ABBO, M.D., J.D., [5] and LISA LEHMANN, M.D., Ph.D.[6]

ABSTRACT

***Background*: Several studies have reported that African Americans are more likely than whites to prefer aggressive treatments at the end of life.**

***Objective*: Since the medical information presented to subjects is frequently complex, we hypothesized that apparent differences in end-of-life preferences and decision making may be due to disparities in health literacy. A video of a patient with advanced dementia may overcome communication barriers associated with low health literacy.**

***Design*: Before and after oral survey.**

***Participants*: Subjects presenting to their primary care doctors.**

***Methods*: Subjects were asked their preferences for end-of-life care after they heard a verbal description of advanced dementia. Subjects then viewed a 2-minute video of a patient with advanced dementia and were asked again about their preferences. For the analysis, preferences were dichotomized into comfort care and aggressive care. Health literacy was measured using the Rapid Estimate of Adult Literacy in Medicine (REALM) and subjects were divided into three literacy categories: low (0–45, sixth grade and below), marginal (46–60, seventh to eighth grade) and adequate (61–66, ninth grade and above). Unadjusted and adjusted logistic regression models were fit using stepwise algorithms to examine factors related to initial preferences before the video.**

***Results*: A total of 80 African Americans and 64 whites completed the interview. In unadjusted analyses, African Americans were more likely than whites to have preferences for aggressive care after the verbal description, odds ratio (OR) 4.8 (95% confidence interval [CI] 2.1–10.9). Subjects with low or marginal health literacy were also more likely than subjects with adequate health literacy to have preferences for aggressive care after the verbal description, OR 17.3 (95% CI 6.0–49.9) and OR 11.3 (95% CI 4.2–30.8) respectively. In adjusted analyses, health literacy (low health literacy: OR 7.1, 95% CI 2.1–24.2; marginal health literacy OR 5.1, 95% CI 1.6–16.3) but not race (OR 1.1, 95% CI 0.3–3.2) was an independent predictor of preferences after the verbal description. After watching a video of advanced dementia, there were no significant differences in the distribution of preferences by race or health literacy.**

***Conclusions*: Health literacy and not race was an independent predictor of end-of-life preferences after hearing a verbal description of advanced dementia. In addition, after viewing a video of a pa-**

[1]General Medicine Unit, Massachusetts General Hospital, Boston, Massachusetts.
[2]Department of Medicine, Boston Medical Center, Boston, Massachusetts.
[3]Harvard Medical School, Boston, Massachusetts.
[4]Brigham and Women's Hospital, Boston, Massachusetts.
[5]Section of General Internal Medicine, Department of Medicine, University of Chicago, Chicago, Illinois.
[6]Boston, Massachusetts.

tient with advanced dementia there were no longer any differences in the distribution of preferences according to race and health literacy. These findings suggest that clinical practice and research relating to end-of-life preferences may need to focus on a patient education model incorporating the use of decision aids such as video to ensure informed decision-making.

INTRODUCTION

Several studies have reported that African Americans are more likely than whites to prefer more aggressive treatments at the end of life.[1–9] Speculation about this finding has focused on African Americans' unique cultural and spiritual belief systems,[1,4–7,9] and their lack of trust in the health care system perhaps due to historical experiences such as Tuskegee, in which African American men were deprived of treatment for syphilis as an experiment.[1–4] Although low health literacy is a prevalent barrier to accurate communication and could confound the relationship between race and informed decision-making, no prior studies on end-of-life decision-making have explored the confounding effects of health literacy.

More than 90 million adults in the United States have insufficient literacy skills to be independent health care users,[10] and low health literacy is associated with race, age, education, and having chronic illness.[11] Distinguishing the effect of health literacy from race is important in this context.[12] Ethnicity and culture are factors that may shape preferences and should be respected. By contrast, low health literacy is an obstacle to communicating information for informed decision-making that we should strive to surmount so that we can accurately elicit and respect individual values.

There are multiple communication techniques that could be used to overcome barriers posed by health literacy. Visual images have been shown to improve communication of information with patients.[13–17] Prior interventions to improve decision-making have included print, audio, video, and Web-based pictorial materials,[15,18–25] with mixed results.[26] The medium of video allows patients to envision health states in a manner not easily captured with verbal communication and can both engage people and efficiently communicate information about the experience of illness.[27]

We hypothesized that after hearing a verbal description of advanced dementia, subjects' preferences for care at the end of life for themselves would be independently predicted by health literacy and not race. Additionally, we hypothesized that a video of a patient with advanced dementia would overcome communication barriers associated with low health literacy.

METHODS

Participants

Patients over the age of 40 scheduled to see a general internist at six study sites were eligible to participate. The age of 40 was chosen since most adults after that age have usually been involved with end-of-life decision making for a loved one. Patients were given a flyer outlining the study at the time of their clinic visit. The study sites consisted of urban and suburban primary care clinics affiliated with two teaching hospitals in the greater Boston area. Subjects were excluded if they had previously had a close relationship with a person with advanced dementia, since they will likely have had first-hand experience and knowledge of the disease. Previous work suggested that subjects who had a close relationship strongly favored comfort care and that a video did not better inform preferences.[27] Subjects were also excluded if, in the judgment of the physician interviewer, they lacked medical decision-making capacity at the time of the interview, or if they did not speak English, the language of our validated tools. Only subjects who self-reported their race as African American or white were included in the analysis since this study attempted to compare the preferences of African Americans and whites. All interviews were conducted by two white, male physicians between December 1, 2005 and January 31, 2007. Approval for the project was granted by the Institutional Review Boards of the affiliated hospitals.

Design

A structured questionnaire was developed for this study. Survey questions were generated from a review of the medical ethics literature, and consultations with medical ethics, palliative care, geriatric, and neurology experts. Early versions of the survey were tested with subjects recruited from primary care clinics.

After obtaining verbal informed consent, interviewers defined advanced dementia in simple language, highlighting functional impairments based on the Functional Assessment Staging (FAST)[28] criteria including inability to communicate understandably with others, inability to ambulate without assistance,

and inability to feed oneself[27] (See Appendix A for verbal description). We then outlined three levels of medical treatments and the goals associated with each level. The first level, life-prolonging care, aims to prolong life at any cost. It includes all medically indicated treatments. The second level, limited care, aims to maintain physical functioning. It includes treatments such as hospitalization, intravenous fluids, antibiotics, and artificial nutrition and hydration, but excludes attempted cardiopulmonary resuscitation and treatments in the intensive care unit. The third level, comfort care, aims to maximize comfort and to relieve pain. It includes oxygen and analgesics but excludes intravenous therapies and hospitalization unless necessary to provide comfort. Subjects were asked which level they preferred in the event they developed advanced dementia. Subjects not able to select a level of treatments were documented as "don't know."

Each subject next viewed a 2-minute video of a white patient with advanced dementia. The video depicts the salient features of advanced dementia. The narrative that accompanies the video is identical to the verbal description used to assess subjects' initial preferences. The design, content, and structure of the video intervention were reviewed for accuracy by three geriatricians and five neurologists, all of whom specialize in the care of patients with dementia. The video was also designed with the close collaboration of caregivers of patients with advanced dementia.[27] The video was available only in English. (The film clip is available online at advancecareplanning videos.com.

The subject was then asked exactly the same questions as previously regarding preferences for level of medical care. Answers to sociodemographic questions (age, race, gender, etc.) were self-reported by the subjects. Health literacy was measured at the end of the interview using the validated Rapid Estimate of Adult Literacy in Medicine tool (REALM).[29] The survey is available upon request.

Statistical analysis

The main outcome measure was preferences for care after hearing a verbal description of advanced dementia. We dichotomized preferences into two groups: comfort care and aggressive care (life-prolonging care, limited care, and do not know). As others have done, we included subjects preferring "do not know" in the aggressive care group since in clinical practice the default for such patients is life-prolonging treatments.[27,30]

Our measure for health literacy was the 66-word REALM.[29] This is a 2- to 3-minute English test of medically relevant vocabulary. The REALM is a valid test of word pronunciation and has been shown to correlate well with tests that evaluate a range of literacy skills.[31] As others have done, we defined three categories for literacy: low literacy (REALM score of 0–45, sixth grade and below); marginal literacy (REALM score of 45–60, seventh to eighth grade); and adequate literacy (REALM score of 61–66, ninth grade and above).[11,32] Fisher's exact test was used to compare the proportion of subjects who chose comfort care across levels of health literacy.

Univariate analyses for subject characteristics (health literacy, age, gender, education, marital status, religion, religious attendance, and health status) with race and preferences for care were conducted utilizing the Fisher's exact test. Unadjusted odds ratios (ORs) were calculated using contingency tables to compare subject demographic characteristics to their preferences.

Multivariate logistic regression analyses were used to identify factors independently associated with preferences. Stepwise algorithms were used to assess the independent role of each variable. Factors significant at $p < 0.10$ were retained. Factors thus selected were simultaneously entered into the logistic regression model. Another backward stepwise algorithm, using $p < 0.05$, was used to determine the final model.

The distribution of levels of care after the verbal description compared to the distribution of preferences after the video were analyzed utilizing the McNemar's test of correlated proportions. All p values are two-tailed. Data were analyzed using SAS software, version 9.1 (SAS Institute, Cary, NC).

RESULTS

Study participants

A total of 214 subjects were approached to participate in the study, of whom 173 (81%) agreed to be interviewed. The most common reason given for not participating was lack of time. Of the 173 subjects recruited for the study, 23 were disqualified due to a prior history of a close relationship with a person with advanced dementia. Four subjects were excluded due to being Asian American (1), Native American (1), or Latino (2); 2 subjects were excluded from the analysis due to refusal to participate in the health literacy assessment. The resulting dataset included 144 subjects.

Table 1 describes the characteristics of the survey sample. Of the 144 subjects, 64 (44%) were

Table 1. Characteristics of Study Sample

Characteristics	*Total*	*African Americans*	*Whites*	p *value*
Number of subjects, no. (%)	144	80 (56)	64 (44)	
Age, mean (SD)	57 (11)	56 (11)	58 (12)	0.3
Women, no. (%)	92 (64)	55 (69)	37 (58)	0.22
Health literacy, no. (%)				<0.0001
Low	27 (19)	23 (29)	4 (6)	
Marginal	30 (21)	26 (33)	4 (6)	
Adequate	87 (60)	31 (39)	56 (88)	
Education, no. (%)				<0.0001
High school or less	80 (56)	65 (81)	15 (23)	
College or beyond	64 (44)	15 (19)	49 (77)	
Marital status, no. (%)				0.001
Married	77 (53)	33 (42)	44 (69)	
Nonmarried	67 (47)	47 (59)	20 (31)	
Religion, no. (%)				<0.0001
Christian	116 (81)	74 (93)	42 (66)	
Non-Christian	28 (19)	6 (8)	22 (34)	
Religious attendance, no. (%)				0.0001
Two times per month of more	61 (42)	43 (54)	18 (28)	
One time per month or less	23 (16)	16 (20)	7 (11)	
Never	60 (41)	21 (26)	39 (61)	
Self-reported health status, no. (%)				0.0002
Very healthy	77 (53)	31 (39)	46 (72)	
Somewhat healthy	61 (42)	44 (55)	17 (27)	
Not healthy	6 (4)	5 (6)	1 (2)	

SD, standard deviation.

whites and 80 (56%) were African Americans. African Americans had significantly lower health literacy, lower education, more religious attendance, and poorer health.

Outcomes

After hearing a verbal description of advanced dementia, treatment preference was associated with race, health literacy, education, and health status (Table 2).

Preferences after verbal description by race

Of the 64 whites, 55 (86%) preferred comfort care and 9 (14%) preferred aggressive care (Fig. 1). Of the 80 African Americans, 45 (56%) preferred comfort care and 35 (44%) preferred aggressive care (Fig. 1). Few whites or African Americans indicated "do not know" as their preferences after the verbal description, 2 (3%) and 5 (6%) subjects, respectively.

Unadjusted analysis revealed that compared to whites, African Americans were more likely to have preferences for aggressive care after the verbal description, OR 4.8 (95% CI 2.1–10.9; Table 2). In the multiple logistic regression stepwise model, the effect of race was no longer significant, OR 1.1 (95% CI 0.3–3.2; Table 3).

Preferences after verbal description by health literacy

Preferences after the verbal description were also strongly associated with health literacy (Fig. 1). Of the 27 subjects with low health literacy, 18 (67%) preferred aggressive care. Of the 30 subjects with marginal health literacy, 17 (57%) chose aggressive care. Of the 87 subjects with adequate health literacy, only 9 (10%) selected aggressive care. There was an increasing preference for comfort care after the verbal description with increasing health literacy that was highly statistically significant ($p < 0.0001$).

Unadjusted analyses revealed that subjects with low health literacy were more likely to have preferences for aggressive care after the verbal description when compared to those subjects with adequate health literacy, OR 17.3 (95% CI 6.0–49.9), as did those with marginal health literacy, OR 11.3 (95% CI 4.2–30.8; Table 2). In the multiple logistic stepwise model, health literacy remained a significant and independent predictor of preferences for care (low literacy OR 7.1, 95% CI 2.1–24.2; marginal literacy 5.1, 95% CI 1.6–16.3; Table 3).

Of all the other characteristics (age, education, gender, marital status, religion, religious attendance, and health status), only education remained in the model

Table 2. Unadjusted Odds Ratios of Initial Preferences for Aggressive Care after Verbal Description

Characteristic	Initial preferences for aggressive care[a] after verbal description	
	Unadjusted OR	*95% CI*
Race		
African American	4.8	2.1–10.9
White	1.0	
Health literacy		
Low	17.3	6.0–49.9
Marginal	11.3	4.2–30.8
Adequate	1.0	
Education		
High school or less	15.0	5.0–45.2
College or beyond	1.0	
Gender		
Male	1.0	0.5–2.1
Female	1.0	
Marital status		
Nonmarried	1.8	0.9–3.7
Married	1.0	
Religion		
Non-Christian	1.8	0.7–4.8
Christian	1.0	
Religious attendance		
≤1 month	1.3	0.5–3.8
≥2 month	1.7	0.8–3.7
Never	1.0	
Health status		
Not healthy	1.6	0.3–9.7
Somewhat healthy	2.1	1.0–4.4
Very healthy	1.0	

[a]Aggressive care includes those subjects who chose one of the following: life-prolonging care, limited care, or don't know.
OR, odds ratio; CI, confidence interval.

after the stepwise approach (Table 3). Education was also associated with preferences after the verbal description. Subjects with a high school education or less were more likely to have preferences for aggressive care when compared to those with a college or higher education, OR 15.0 (95% CI 5.0–45.2; Table 2). In adjusted analysis controlling for race and health literacy, education was still a significant and independent predictor of preferences after the verbal description, OR 4.5 (95% CI 1.1–18.6; Table 3).

Change in initial preferences after the video

The distribution of subjects' preferences after viewing the video changed significantly compared to the distribution of preferences after hearing the verbal description ($p < 0.0001$). The vast majority of subjects across both races and all health literacy groups chose comfort care after viewing the video (Fig. 2). After the video, no whites and 10 (13%) African Americans preferred aggressive care (Fig. 2). After the video, 3 (11%) subjects with low health literacy and 3 (10%) subjects with marginal health literacy preferred aggressive care. Of the subjects with adequate health literacy, 4 (3%) preferred aggressive care after the video.

DISCUSSION

Previous studies have suggested that African Americans are more likely than whites to prefer aggressive care at the end of life. This finding was observed in the current study as well; however, adjusted analyses reveal that health literacy mediates the relationship between race and end-of-life preferences for African Americans. The influence of health literacy on end-of-life decision-making reflected a dose-response effect with increasing health literacy leading toward comfort care. This suggests that race should not be the primary focus of research and patient care regarding end-of-life preferences. Rather, health literacy and a heightened commitment to patient education and communication should be the top priority. The importance of patient education and communication is corroborated by the finding in our study that a video intervention designed to compensate for low health literacy skills had a large impact on patient preferences. Indeed, after viewing the video there were no significant differences in end-of-life preferences by race or health literacy status. Carefully crafted video decision aids designed to overcome barriers posed by health literacy may improve decision-making at the end of life for patients with low health literacy. Furthermore, racial disparities in features of clinical medicine that may be heavily dependent on health literacy, such as we found to be the case in this study, can be influenced by interventions like video that target deficits in health literacy. Video offers a unique and widely accessible means of communication that engages patients in a way not achieved with words.

None of the studies previously cited regarding African American preferences at the end of life studied health literacy,[1–9] and few asked about education.[2,3,6] Education is closely related to health literacy, but the two variables clearly represent different constructs. Education level represents how far a person progressed in formal education, not how much literacy skill he/she acquired. Various forces such as social promotion can lead to significant variance between education level and observed literacy. In this country, the average adult English reading level is between eighth and ninth grade,[33] and, the quality of educa-

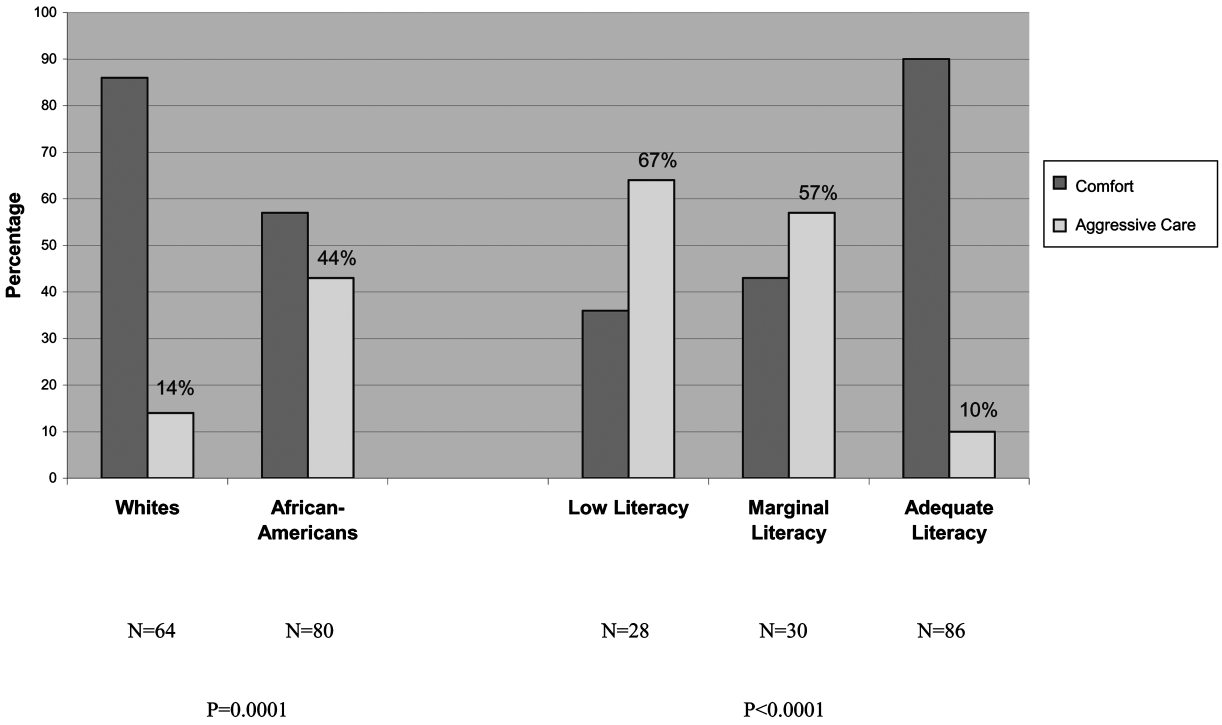


FIG. 1. Preferences after verbal description by race and health literacy. The Fisher's exact test was used to compare each variable (race, literacy) with the proportion of subjects who chose comfort care.

tional outcomes varies significantly.[34] Our results show that both education and health literacy were independently associated with end-of-life care preferences. We posit that health literacy will be more closely related to communication barriers than education because health literacy is a direct measure of a patient's current skills. Future studies exploring in detail the increasingly recognized roles of health literacy and education at the end of life are important.[35–36]

Our study has several important limitations. First, the distribution of low or marginal health literacy between African Americans and whites was not parallel. There were only 8 whites with low or marginal health literacy. Future research on the relationship between race, health literacy, and end-of-life preferences could benefit from a more balanced or simply larger analytic sample. This would permit exploration of additional potential confounders, such as socioeconomic status. In addition, our sample was drawn from the metro Boston area. These findings should be evaluated in a nationally representative sample that includes other large minority groups such as Latinos and Asian Americans.

We used a before and after study design in which individual subjects heard the description of dementia twice, verbally and then with the video. Subjects with limited health literacy may simply have benefited from the repetition of information. Future studies isolating the effect of the video would be helpful. Additionally, preferences for level of care at the end of life may change over the course of time. Our study looked at the preferences of subjects at a particular point in time.

Table 3. Adjusted Odds Ratios for Initial Preferences for Aggressive Care after Verbal Description

	Initial preferences for aggressive care after verbal description	
Characteristic[a]	*Unadjusted OR*	*95% CI*
Race		
African American	1.1	0.3–3.2
White	1.0	
Health literacy		
Low	7.1	2.1–24.2
Marginal	5.1	1.6–16.3
Adequate	1.0	
Education		
≤High school	4.5	1.1–18.6
≥College	1.0	

[a]Other characteristics (age, gender, religion, religious attendance and health status) were excluded from the stepwise model as they were not related to the outcome at a $p < 0.10$.

OR, odds ratio; CI, confidence interval.

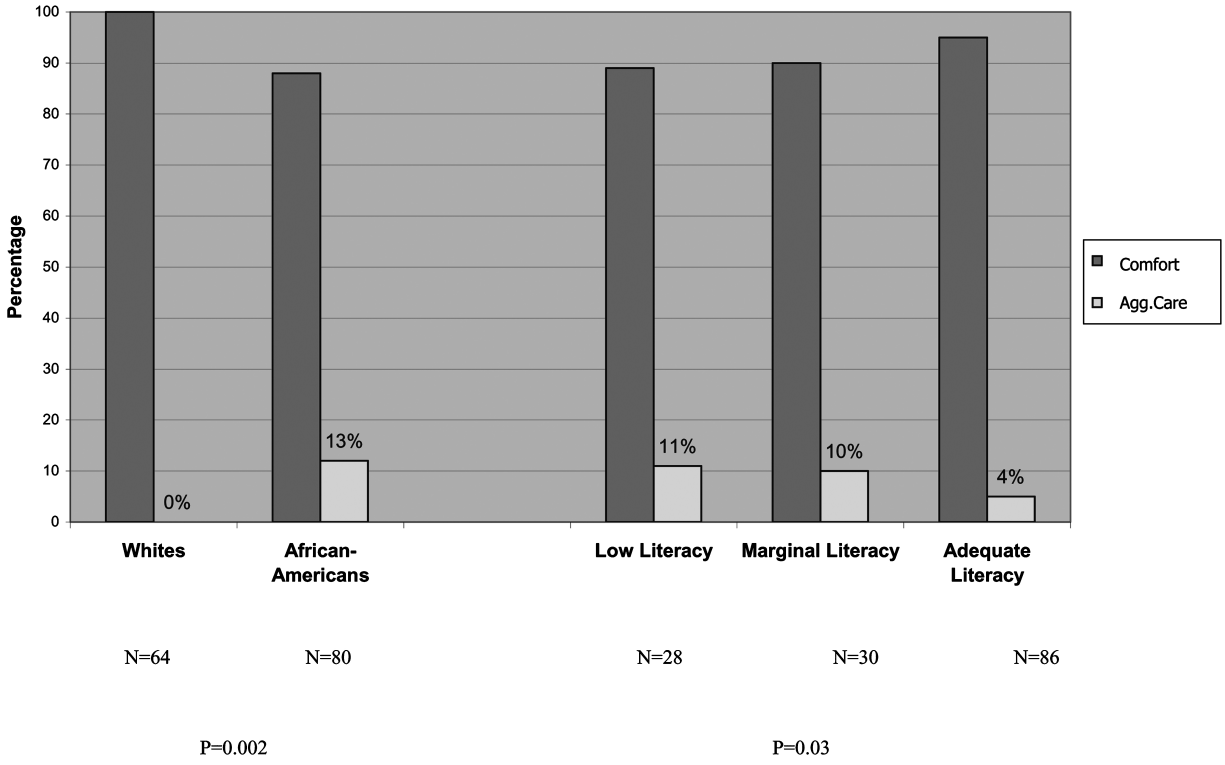


FIG. 2. Preferences after video by race and health literacy. The Fisher's exact test was used to compare each variable (race, literacy) with the proportion of subjects who chose comfort care.

Future work looking at the stability of preferences over time would be helpful, especially in the context of completing advance directives or more formalized advance care planning that incorporates the video.

Video is a powerful medium that can be manipulated to sway patient perspectives. We studied the use of a single video clip that was designed to portray advanced dementia in order to overcome barriers posed by health literacy. Future studies exploring the designing and filming of videos to overcome literacy barriers would be useful. It would also be interesting to explore the preferences of subjects using other video clips that varied the features of the patient, such as race and gender. It is important to note that special care needs to be taken to develop patient education materials and decision-aids that empower—not manipulate—patients. Our study asked questions about preferences if patients would be in a state of advanced dementia, a common end-of-life scenario. Preferences for end-of-life care in other disease states may be different. Other studies using disease states such as a persistent vegetative state, severe stroke, end-stage chronic obstructive pulmonary disease (COPD) and severe congestive heart failure (CHF) may be fruitful.

Discussions regarding treatments at the end of life will increasingly involve a more diverse pool of patients, and consideration of each patient's culture is important. Yet our data show that accepting a patient's preferences based on a verbal conversation about a future disease state may not reflect a patient's fully informed preferences for end-of-life treatments: the diversity of preferences that have often been attributed to racial and cultural differences are a result of differences in health literacy, not race. Video in addition to improved verbal communication may be one means to overcome barriers posed by health literacy. Prematurely accepting aggressive preferences for end-of-life treatments may inadvertently condemn patients to an unwanted and misunderstood course of medical treatments at the end of life.

ACKNOWLEDGMENTS

We thank Drs. Michael Barry, Susan Mitchell, and Miguel Ariza for their contribution to the project.

Dr. Volandes 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. Garrett JM, Harris RP, Norburn JK, Patrick DL, Danis M: Life-sustaining treatments during terminal illness: Who wants what? J Gen Intern Med 1993;8:361–368.
2. Caralis PV, Davis B, Wright K, Marcial E: The influence of ethnicity and race on attitudes toward advance directives, life-prolonging treatments, and euthanasia. J Clin Ethics 1993;4:155–165.
3. O'Brien LA, Grisso JA, Maislin G, LaPann K, Krotki KP, Greco PJ, Siegert EA, Evans LK: Nursing home residents' preferences for life-sustaining treatments. JAMA 1995; 274:1775–1779.
4. Blackhall LJ, Frank G, Murphy ST, Michel V, Palmer JM, Azen SP: Ethnicity and attitudes towards life sustaining technology. Soc Sci Med 1999;48:1779–1789.
5. Crawley L, Payne R, Bolden J, Payne T, Washington P, Williams S: Palliative and end-of-life care in the African American community. JAMA 2000;284:2518–2521.
6. Hopp FP, Duffy SA: Racial variations in end-of-life care. J Am Geriatr Soc 2000;48:658–663.
7. Johnson KS, Elbert-Avila KI, Tulsky JA: The influence of spiritual beliefs and practices on the treatment preferences of African Americans: A review of the literature. J Am Geriatr Soc 2005;53:711–719.
8. Sehgal A, Galbraith A, Chesney M, Schoenfeld P, Charles G, Lo B: How strictly do dialysis patients want their advance directives followed? JAMA 1992;267:59–63.
9. Klessig J: The effect of values and culture on life-support decisions. West J Med 1992;157:316–322.
10. U.S. Department of Education, National Center for Education Statistics: 2003 National Assessment of Adult Literacy. ⟨http://nces.ed.gov/NAAL/index.asp?file=AssessmentOf/HealthLiteracy.asp&PageID=1 61⟩ (Last accessed July 1, 2007).
11. Paasche-Orlow MK, Parker RM, Gazmararian JA, Nielsen-Bohlman LT, Rudd RR: The prevalence of limited health literacy. J Gen Intern Med 2005;20:175–184.
12. Paasche-Orlow M: The ethics of cultural competence. Acad Med 2004;79:347–350.
13. Frosch DL, Kaplan RM, Felitti VJ: A randomized controlled trial comparing internet and video to facilitate patient education for men considering the prostate specific antigen test. J Gen Intern Med 2003;18:781–787.
14. Morgan MW, Deber RB, Llewellyn-Thomas HA, Gladstone P, Cusimano RJ, O'Rourke K, Tomlinson G, Detsky AS: Randomized, controlled trial of an interactive videodisc decision aid for patients with ischemic heart disease. J Gen Intern Med 2000;15:685–693.
15. Edwards A, Elwyn G, Mulley A: Explaining risks: Turning numerical data into meaningful pictures. BMJ 2002; 324:827–830.
16. Houts PS, Doak CC, Doak LG, Loscalzo MJ: The role of pictures in improving health communication: A review of research on attention, comprehension, recall, and adherence. Patient Educ Couns 2006;61:173–190.
17. Ditto PH, Danks JH, Smucker WD, Bookwala J, Coppola KM, Dresser R, Fagerlin A, Gready RM, Houts RM, Lockhart LK, Zyzanski S: Advance directives as acts of communication: A randomized controlled trial. Arch Intern Med 2001;161:421–430.
18. Woolf SH, Chan EC, Harris R, Sheridan SL, Braddock CH 3rd, Kaplan RM, Krist A, O'Connor AM, Tunis S: Promoting informed choice: Transforming health care to dispense knowledge for decision making. Ann Intern Med 2005;143:293–300.
19. Mitchell SL, Tetroe J, O'Connor AM: A decision aid for long-term tube feeding in cognitively impaired older persons. J Am Geriatr Soc 2001;49:313–316.
20. O'Connor AM, Stacey D, Rovner D, Llewellyn-Thomas H, Rovner D, Holmes-Rovner M, Tait V, Tetroe J, Fiset V, Barry M, Jones J: Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2001(3):CD001431.
21. Ouslander JG, Tymchuk AJ, Rahbar B: Health care decisions among elderly long-term care residents and their potential proxies. Arch Intern Med 1989;149:1367–1372.
22. Deyo RA, Cherkin DC, Weinstein J, Howe J, Ciol M, Mulley AG, Jr: Involving patients in clinical decisions: Impact of an interactive video program on use of back surgery. Med Care 2000;38:959–969.
23. Dunn RA, Shenouda PE, Martin DR, Schultz AJ: Videotape increases parent knowledge about poliovirus vaccines and choices of polio vaccination schedules. Pediatrics 1998;102:e26.
24. Wallace C, Leask J, Trevena LJ: Effects of a web based decision aid on parental attitudes to MMR vaccination: A before and after study. BMJ 2006;332:146–149.
25. Barry MJ: Health decision aids to facilitate shared decision making in office practice. Ann Intern Med 2002;136:127–135.
26. Flory J, Emanuel E: Interventions to improve research participants' understanding in informed consent for research: A systematic review. JAMA 2004;292:1593–1601.
27. Volandes A, Lehmann LS, Cook EF, Shaykevich S, Abbo ED, Gillick MR: Using video images of advanced dementia in advance care planning. Arch Intern Med 2007;167: 828–833.
28. Sclan SG, Reisberg B: Functional assessment staging (FAST) in Alzheimer's disease: reliability, validity, and ordinality. Int Psychogeriatr 1992;4(Suppl 1):55–69.
29. Doak CC, Doak LG, Root JH: *Teaching Patients with Low Literacy Skills, 2nd ed.* Philadelphia: J.B. Lippincott, 1996.
30. Goodlin SJ, Zhong Z, Lynn J, Teno JM, Fago JP, Desbiens N, Connors AF Jr, Wenger NS, Phillips RS: Factors associated with use of cardiopulmonary resuscitation in seriously ill hospitalized adults. JAMA 1999;282:2333–2339.
31. Davis TC, Long SW, Jackson RH, Mayeaux EJ, George RB, Murphy PW, Crouch MA: Rapid estimate of adult literacy in medicine: A shortened screening instrument. Fam Med 1993;25:391–395.
32. Davis TC, Wolf MS, Bass PF, 3rd, Thompson JA, Tilson HH, Neuberger M, Parker RM: Literacy and misunderstanding prescription drug labels. Ann Intern Med 2006; 145:887–894.
33. Communicating with patients who have limited literacy skills. Report of the National Work Group on Literacy and Health. J Fam Pract 1998;46:168–176.
34. National Assessment of Adult Literacy (NAAL): U.S. Department of Education. ⟨http://nces.ed.gov/NAAL/index.asp⟩ (Last accessed July 1, 2007).

35. White DB, Braddock CH, 3rd, Bereknyei S, Curtis JR: Toward shared decision making at the end of life in intensive care units: Opportunities for improvement. Arch Intern Med 2007;167:461–467.
36. Campbell MJ, Edwards MJ, Ward KS, Weatherby N: Developing a parsimonious model for predicting completion of advance directives. J Nurs Scholarsh 2007;39:165–171.

Address reprint requests to:
Angelo E. Volandes, M.D.
General Medicine Unit
Massachusetts General Hospital
50 Staniford Street, 9th Floor
Boston, MA 02114

E-mail: avolandes@partners.org

Appendix A. Narrative Describing Advanced Dementia

"I am going to describe to you an illness called advanced dementia, like advanced Alzheimer's dementia, that you may or may not be familiar with. Advanced dementia is an incurable disease of the brain in which one is not able to communicate with others. People in advanced dementia are not able to move around or walk, get out of bed independently, eat by oneself, or communicate understandably with others. People with advanced dementia often have difficulty chewing or swallowing, and require assistance with feeding oneself. Advanced dementia is an incurable disease and most commonly occurs after many years of Alzheimer's disease or as the result of strokes. People are not able to answer any questions or tell you about themselves."

Narrative Describing the Goals of Care

"I am going to ask you a question about your preferences for medical care if you had a disease called advanced dementia. I will ask you what you prefer. You have three choices for medical care if you had this condition. I will first review these three choices with you. The three choices for medical care that i want you to think about for advanced dementia are life-prolonging care, limited care, and comfort care.

Life-prolonging care

The goal of this category of care is to prolong life. There are no limits to care. This choice includes everything a modern hospital has to offer to maintain your life. Such procedures include: cardiopulmonary resuscitation or CPR in which a doctor pushes on your chest when the heart stops and will often use electricity to shock the heart. Being placed on a breathing machine, also known as life support, in which a tube is placed down your throat into the lungs. And other medical procedures performed in the intensive care unit or ICU. The goal is to prolong life.

Limited care

The goal of this category is to maintain physical and mental functions. Care will depend on your physical and mental functioning. Such care includes intravenous (IV) therapies like antibiotics, feeding tubes and hospitalization. But does not include CPR and ICU care. The goal is to maintain physical and mental functioning.

Comfort care

The goal of this category is to maximize comfort. Only measures that comfort or relieve pain are performed. The aim is to relieve pain and to be kept as pain-free as possible. Comfort care does not include CPR respirators, ICU care, and generally would not include IV therapy, feeding tubes, or hospitalization. The goal is maximizing comfort and relieving pain.

Literacy, Social Stigma, and HIV Medication Adherence

Katherine R. Waite, BA[1], Michael Paasche-Orlow, MD MA MPH[2], Lance S. Rintamaki, PhD[3,4], Terry C. Davis, PhD[5], and Michael S. Wolf, PhD MPH[1]

[1]Health Literacy and Learning Program, Institute for Healthcare Studies, Division of General Internal Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA; [2]Section of General Internal Medicine, Department of Medicine, Boston University School of Medicine, Boston, MA, USA; [3]Department of Communication, State University of New York, Buffalo, NY, USA; [4]Department of Health Behavior, State University of New York, Buffalo, NY, USA; [5]Louisiana State University Health Sciences Center, Shreveport, LA, USA.

BACKGROUND: Prior studies have linked limited literacy to poorer HIV medication adherence, although the precise causal pathways of this relationship have only been initially investigated.

OBJECTIVE: To examine whether social stigma is a possible mediator to the relationship between literacy and self-reported HIV medication adherence.

DESIGN: Structured patient interviews with a literacy assessment, supplemented by medical chart review, were conducted among patients receiving care at infectious disease clinics in Shreveport, Louisiana and Chicago, Illinois. Literacy was measured using the Rapid Estimate of Adult Literacy in Medicine (REALM), while stigma was measured using items taken from the Patient Medication Adherence Questionnaire (PMAQ).

PARTICIPANTS: Two hundred and four consecutive patients participated.

RESULTS: Approximately one-third of the patients (30.4%) were less than 100% adherent to their regimen, and 31.4% had marginal (7th–8th grade) or low (≤ 6th grade) literacy. In multivariate analyses, patients with low literacy were 3.3 times more likely to be non-adherent to antiretroviral regimens (95% CI 1.3–8.7; p <0.001). Perceived social stigma was found to mediate the relationship between literacy and medication adherence (AOR 3.1, 95% CI 1.3–7.7).

CONCLUSIONS: While low literacy was a significant risk factor for improper adherence to HIV medication regimens in our study, perceived social stigma mediated this relationship. Low literacy HIV intervention strategies may also need to incorporate more comprehensive psychosocial approaches to overcome stigma barriers.

KEY WORDS: literacy; stigma; HIV; medication; adherence.
J Gen Intern Med 23(9):1367–72
DOI: 10.1007/s11606-008-0662-5

Received October 22, 2007
Revised March 27, 2008
Accepted April 28, 2008
Published online June 19, 2008

INTRODUCTION

Several studies have identified limited literacy as a risk factor for poor HIV medication adherence, despite evidence that questions this association.[1–4] Conceptually, whether this relationship is construed as intuitive is controversial, as the actual causal pathways linking literacy to health outcomes remain unclear.[5–7] Nonetheless, it is important to understand the nature of the literacy-adherence relationship and what specific factors might plausibly explain it.

A possible mediating variable that might partially explain the association between literacy and medication adherence, specifically in the context of HIV/AIDS, could be concern for social stigma. Stigma has been labeled as the most formidable social and psychological aspect of the HIV experience.[8,9] HIV stigma includes "prejudice, discounting, discrediting, and discrimination directed at people perceived to have AIDS or HIV, and the individuals, groups, and communities with which they are associated."[10] Large segments of the public remain uneducated about HIV and how it is transmitted, which promotes fear and antipathy toward those infected with the virus.[11] These sentiments may often translate into biased and discriminatory actions. Concern for such stigma is widely reported among people living with HIV, which subsequently affects their social, psychological, and physical well-being.[12,13] Concern for and experience with stigma has been linked to elevated stress, depression, impaired immune response, and high suicide rates among those living with HIV.[14–20] It can also lead people to hide their serostatus from others, and avoid healthcare or forego their antiretroviral medications.[21–23] Specifically, patients may miss doses of their regimen if their schedule requires them to take the medicine at inopportune times and in public environments.[22,24,25] Rintamaki and colleagues found that a greater concern for social stigma was a significant independent predictor of HIV medication adherence.[23]

It is possible that patients with limited literacy skills may be more sensitive to matters of shame and stigma as a result of their co-existing concern for social stigma related to their limited reading proficiency, among other psychosocial issues.[2,26–29] In earlier studies, HIV-infected patients with inadequate health literacy were found to be significantly more likely to report negative health care perceptions and experiences, and to be less confident in their ability to self manage their disease than those with adequate health literacy. [2,30]

No study to our knowledge has directly investigated the relationship between concern for social stigma and literacy level or, specifically, whether perceived stigma mediates the relationship between literacy and medication adherence. We

sought to investigate the relationship between literacy, social stigma concerns and HIV medication adherence among a diverse cohort of patients.

METHODS

Sample

The study sample and methods have been previously described in detail.[2,31,32] From June to September 2001, we enrolled a total of 204 consecutive HIV-infected patients receiving medical care who were prescribed one or more antiretroviral medications and received medical care through outpatient infectious disease clinics at the Northwestern Memorial Hospital (Chicago site) and the Louisiana State University Health Sciences Center at Shreveport (LSUHSC). Patients who were on their regimen for less than two weeks were excluded from participation, as well as those having any of the following conditions: (1) dementia; (2) blindness or severely impaired vision not correctable with eyeglasses; (3) deafness or hearing problems uncorrectable with a hearing aid; (4) too ill to participate in the survey. Approval for human subjects research was obtained from institutional review boards at both study sites prior to consenting patients to the study.

Data and Procedure

Trained research assistants received referrals of interested and eligible patients from clinic health providers, gathered informed consent, and conducted a structured interview with recruited patients in a private room. Information collected at this time included literacy level, stigma concerns, and HIV medication adherence.

Medication Adherence

Patients were given the low literacy accessible version of the Patient Medication Adherence Questionnaire (PMAQ) and asked to identify the medications in their current regimen, as well as self-report any recent missed doses using pages that contained names and color photographs of common HIV medications.[33,34] This version of the PMAQ had been revised to simplify the test by limiting the recall period for missed doses and adding visual cues to aid in medication identification. Information on patient antiretroviral agents, comorbidities, and non-HIV prescriptions was obtained through medical chart reviews. Four questions were asked pertaining to any missed doses over the past four days. Proper adherence was determined if patients reported no skipped doses over the past four days, whereas non-adherence was determined if any doses in the patient's regimen were missed during the past four day interval.[1,2,23,31,32]

Measurement

HIV-Related Social Stigma. Patients' concern with HIV-related stigma was measured using three items from the PMAQ, which is a 25-item scale that assesses psychosocial barriers to adherence.[33] Items included statements related to problems associated with adhering to HAART regimens. Three statements were presented to patients that addressed internalized stigma issues: "I am embarrassed to get my medicines from a drug store;" "I don't want people to see me take my HIV medicines;" and "Taking my medicines reminds me that I have HIV." Patients were asked to respond to each statement by endorsing the item on a three-point scale (agree, not sure, disagree).

Conceptually, the first two statements speak to participants' sensitivity to exposing their medications and, potentially, their HIV status. The dilemmas these statements represent have been detailed in earlier work regarding HIV patients' stigma concerns for being seen using antiretroviral medications or accessing HIV-related services.[21,22] The first and third statements also speak to people who may have internalized HIV stigma, whose embarrassment and heightened identity salience around their medications is hypothesized to correlate with high concerns for HIV stigma and discrimination.[35–38]

A total score of social stigma concern was calculated (range 3 to 9), and perceived HIV-related stigma concern was categorized as either low (3–5), moderate (6–7), or high (8–9). These items have previously been found to demonstrate high internal consistency (> 0.85), loading on a single stigma factor (Eigenvalue=1.60; factor loadings 0.73, 0.76, and 0.80, respectively).[23]

Literacy. Patient literacy was assessed using the Rapid Estimate of Adult Literacy in Medicine (REALM).[39] For this literacy assessment, patients are asked to read aloud 66 medical terms while a trained research assistant scores the REALM based on number of words pronounced correctly. Classification of literacy are as follows: third grade or less (0–18), fourth to sixth grade (19–44), seventh to eighth grade (45–60), and ninth grade and above (61–66). In health care studies where patients need only be categorized as low (scores 0–44), marginal (scores 45–60) or adequate (scores 61–66) readers, the information provided by the REALM is generally sufficient. The REALM is highly correlated with standardized reading tests and the Test of Functional Health Literacy in Adults (TOFHLA).[40,41]

Analysis Plan

Statistical Analysis. Chi-square and student's t-tests were used to evaluate the association between patient literacy, demographic (age, gender, race, insurance coverage, employment, monthly income, site) and clinical (number of HIV and non-HIV medications currently taken, comorbidity, treatment in the past 6 months for mental illness or illicit drug use) characteristics, social stigma concerns, and self-reported adherence to HAART regimens (100% adherence vs. < 100% adherence, past four days). Patient literacy was classified either as low (6th grade and below), marginal (7th–8th grade) or adequate (9th grade and higher). Multivariate logistic regression models were used to estimate the independent relationship between low literacy and the outcomes of associated concerns of social stigma and medication adherence while controlling for potential confounding variables (age, gender, race, site) and additional risk factors for medication non-adherence (number of HIV medications in regimen, other medications taken, comorbidity, history of mental illness and/or illicit drug use).

Mediational Analysis. We used mediational analysis to analyze the pathways linking literacy and HIV medication adherence.[42] Mediating variables are those thought to lie in a causal pathway between the main predictor variable and the outcome. The independent relationship between literacy and medication adherence was revisited, adjusting for all exogenous covariates and potential interaction effects (baseline model). Next, the relationship between literacy and social stigma concerns were examined. Finally, social stigma was added to the baseline model as a mediator, and changes in odds ratios for patient literacy were analyzed. Model calibration and discrimination was estimated using the Hosmer-Lemeshow goodness-of-fit chi-square test and the c-statistic from receiver operating characteristic (ROC) curves. All statistical analyses were performed using STATA, version 8.0 (College Station, TX).

RESULTS

Sample Characteristics

The mean age of patients was 40.1 years, 45.1% were African American and 79.9% were male. Over half of respondents (55.9%) were unemployed, 39.7% had a household income of less than $800/month, and 27.5% did not carry any health insurance. More than 60% of patients reported at least some college education. Approximately one-third of patients had limited literacy skills; 11.3% were reading at or below a sixth grade level (low literacy) and 20.1% were reading at a seventh to eighth grade level (marginal literacy). More than half (52.5%) of patients were also being treated for a non-HIV related chronic illness. Nearly one-third reported receiving mental health services and 9.3% received treatment for alcohol or illicit drug use in the past six months. Significant differences in demographic and clinical characteristics were noted across literacy levels (see Table 1). Patients with low literacy were more likely to be African American, lower educated, male, employed but uninsured, and from the Shreveport site.

Table 1. Characteristics of Sample, Stratified by Literacy Level

Variable	Literacy Level			P value
	Adequate (n=140)	Marginal (n=41)	Low (n=23)	
Age				0.91
< 40	57.9	63.4	56.5	
40–50	29.3	26.8	26.1	
> 50	12.8	9.8	17.4	
Gender				0.03
Male	78.3	65.9	84.3	
Race				<0.001
African American	31.4	68.3	86.9	
Education				<0.001
< High school	5.7	22.0	34.8	
High school graduate	17.9	43.9	43.5	
> High school	76.4	34.1	21.7	
Monthly income				.06
< $800	33.6	43.9	69.6	
$800–$999	24.3	24.4	13.0	
$1000–$1500	11.4	9.8	0.0	
> $1500	30.7	21.9	17.4	
Employment				<0.001
Unemployed	73.9	56.1	52.9	
Employed, part-time	13.0	17.1	15.0	
Employed, full-time	13.1	26.8	32.1	
Insurance				<0.001
Private	33.6	21.9	0.0	
Medicare	20.0	22.0	13.0	
Medicaid/free care	46.4	56.1	87.0	
Site				0.02
Shreveport	50.7	43.9	78.3	
Chicago	49.3	56.1	21.7	
Social stigma concerns				0.007
Low	43.6	56.1	21.7	
Moderate	38.6	26.8	30.4	
High	17.9	17.1	47.9	
No. of HIV medications in regimen				0.17
1– 2	25.9	35.5	45.0	
3 or more	74.1	64.5	55.0	
Non-adherence in past 4 days				0.01
1 or more missed doses	30.0	19.5	52.2	

Social Stigma Concern and Medication Adherence

Over 70% of the patients were taking three or more antiretroviral medications in addition to a mean of 3 (SD=2.9) non-HIV prescription medications. Patients with low literacy had the highest reported rate of non-adherence (52.2%) and individuals with marginal literacy skills were the least likely to self-report missing any doses of antiretroviral medications (19.5%). Patients who reported moderate or high levels of social stigma concern were also more likely to be non-adherent compared to those with low levels of social stigma concern (high – 46.4%, moderate – 30.6%, and low – 22.5%, p=0.01).

Multiple logistic regression models that included social stigma concerns and medication adherence as dependent variables were analyzed using generalized estimating equations (GEE) for binomial data (Table 2). Low literacy (≤6th grade) was a significant independent predictor of high concern for social stigma (Adjusted Odds Ratio (AOR) 3.1, 95% confidence interval (CI) 1.8–9.7), and medication non-adherence in the past four days (AOR 3.3, 95% CI 1.3–8.7; Table 2). Marginal literacy (7th–8th grade) was not significantly associated with social stigma concern or adherence. Multiple regression analyses were then repeated to examine the relationship between social stigma and HIV medication adherence, without literacy in the model. A high level of social stigma concern was also found to be a significant independent predictor of medication non-adherence in the past four days (AOR 3.7, 95% CI 1.5–9.1; Model 2, Table 3).

Mediational Analyses

The multivariate model for medication non-adherence was repeated in mediational analyses, including the hypothesized potential mediating factor of social stigma concern (Table 3). After concern for social stigma was entered into the model the relationship between literacy and adherence attenuated to a point of non-significance (AOR 2.1, 95% CI 0.7–6.5). High concern for social stigma was a significant independent

Table 2. Adjusted Odds Ratios (OR) for Social Stigma Concern and Medication Non-Adherence, by Literacy Level

Outcome	Literacy Level		
	Adequate (n=140)	Marginal (n=41)	Low (n=23)
High social stigma concern, %	17.9	17.1	47.9
Crude OR (95% CI)	1.0	0.9 (0.4–2.4)	4.2 (1.7–10.6)
Adjusted OR (95% CI) *[a]	1.0	0.7 (0.1–6.3)	3.1 (1.8–9.7)
HIV medication adherence, %	70.0	80.5	47.8
Crude OR (95% CI)	1.0	0.5 (0.2–1.2)	2.9 (1.3–6.5)
Adjusted OR (95% CI)	1.0	2.1 (0.8–5.5)	3.3 (1.3–8.7)

Odds ratios adjusted for age, gender, site, insurance coverage, employment status, number of medications in HIV regimen, number of non-HIV prescription medications currently taken, presence of a comorbid chronic condition, treatment for a mental health condition in past six months, and treatment for alcohol or drug use in past six months.
CI – confidence interval

predictor of medication non-adherence in the final model (AOR 3.1, 95% CI 1.3–7.7). Interactions between literacy and social stigma concern were entered into the model; these were not found to be statistically significant.

DISCUSSION

In our study, we re-examined the relationship between literacy and adherence to HIV antiretroviral medications among a diverse sample of patients from two distinct regions of the United States. Approximately one-third of patients in our sample had missed one or more doses in their HAART regimen within the past four days, and low literacy was associated with more than a threefold greater likelihood of missed doses. In mediational analyses, the effect of literacy on medication adherence was reduced by nearly 40% after social stigma concern was included in the model. To our knowledge, this is the first study that documents the association between limited literacy, stigma, and medication adherence. As such, our research advances the national agenda to describe, in more detail, the likely causal pathways linking literacy to health outcomes.

Specifically, higher perceived social stigma mediated the relationship between limited literacy and worse antiretroviral medication adherence. This is not surprising, as it is plausible that stigma concerns might interfere with appropriate processing and understanding of health information, as well as sustained medication-taking behaviors that frequently occur in social situations. Yet stigma concern might impact adherence on a broader level, such as general acceptance of one's condition and the need for antiretroviral therapies.[43] Our findings suggest that patients with low literacy skills are more sensitive to these concerns. Unfortunately we did not assess patients' sense of shame related to domains other than HIV medication-taking behavior, as prior research has shown that patients with low literacy harbor a great amount of shame and stigma relating to their limited literacy skills.

Additional study limitations should be mentioned. We assessed adherence via self-report rather than more objective measures, such as random pill counts, medication event monitoring system (MEMS) caps, or pharmacokinetic laboratory assessments. Although we utilized an existing, validated assessment tool to measure HIV medication adherence, patients may under-report missed doses through questionnaires.[33,34] In addition, our measure of concern for HIV-related stigma was a short, three-item scale derived from a previously validated instrument, which may lack the sensitivity of more in-depth stigma questionnaires. Further psychometric evaluation should be performed to determine the utility of this scale compared to other available tools that may not be so parsimonious and practical for use in clinical settings.[44] Our data is also derived from a cohort of HIV-infected patients interviewed five years ago, and may not directly reflect the experience of those currently on HAART regimens. While more recent advances offer the potential for simplified and less restrictive dosing schedules, adherence still remains a significant challenge for patients with the disease.[45,46] However, our study is one of the first to quantitatively report on the relationship between social stigma and HIV medication adherence among a sample of patients from both urban and rural settings. Therefore, we believe our findings to still be relevant in the present day. Finally, the relatively small sample size and modest number of individuals with low literacy skills further limits the generalizability of study findings.

Despite the limitations, this study suggests that high concern for social stigma is a significant independent predictor of poor medication adherence for those with low literacy. This finding is important, given that most low literacy strategies have focused mostly on simplifying health information without addressing the social circumstances of health care.[47–50] Interventions are needed that extend beyond the 'plain language' programs that have been developed. These methods have resulted in variable success and usually do not lead to improvements in health behaviors.[48,49] While improving health information is imperative, more comprehensive approaches

Table 3. Adjusted Odds Ratios (AOR) for Non-Adherence to HIV Medication Regimen, Past Four Days

Variable	Non-adherence to HAART Regimen		
	Model 1	Model 2	Model 3
	AOR (95% CI)	AOR (95% CI)	AOR (95% CI)
Literacy level			
Adequate (≥ 9th grade)	1.0		1.0
Marginal (7th–8th grade)	2.1 (0.8, 5.5)		0.7 (0.2, 1.8)
Low (≤ 6th grade)	3.3 (1.3, 8.7)		2.1 (0.7, 6.5)
Stigma concerns			
Low		1.0	1.0
Moderate		1.9 (0.8, 4.5)	1.4 (0.6, 3.2)
High		3.7 (1.5, 9.1)	3.1 (1.3, 7.7)

Odds ratios adjusted for age, gender, site, insurance coverage, employment status, number of medications in HIV regimen, number of non-HIV prescription medications currently taken, presence of a comorbid chronic condition, treatment for a mental health condition in past six months, and treatment for alcohol or drug use in past six months.
CI – confidence interval

should address psychosocial concerns by assessing not only patients' understanding, but also their self-efficacy and ability to perform the tasks necessary to promote health.[51] A prior study by Wolf and colleagues already underscores the value of self-efficacy in overcoming literacy barriers in HIV medication adherence.[2] Yet chronic disease self-management programs infrequently mention social stigma specifically as a potential barrier.[51–53] It is possible that social stigma concerns could be addressed through greater patient activation, but this requires further study. Limited evidence is currently available that demonstrates ways to remediate stigma concerns, although interventions targeting patient coping skills and self-efficacy have been cited.[54,55] These types of programs require further testing for use among lower literate HIV patients specifically.

Providers should also be included in the design of effective responses to address stigma and adherence issues for patients with low literacy. Communication skills training modules have already been developed that can aid physicians and other health professionals in discussing the specific implementation of medication schedules within patients' lifestyle and daily routine.[56,57] This allows for preventive problem-solving around potentially difficult scenarios that might lead to missed doses. Ultimately, stigma and literacy concerns should be addressed with strategies that target both patients and providers and have been tested in diverse clinical settings, including community health centers that serve low-income and minority patients who are at greater risk for literacy barriers and may have more challenging social environments.

Acknowledgments: Grant support: *None*

Conflict of Interest: *None disclosed.*

Corresponding Author: *Katherine R. Waite, BA; Health Literacy and Learning Program, Institute for Healthcare Studies, Division of General Internal Medicine, Feinberg School of Medicine, Northwestern University, 750 N. Lake Shore Drive, 10th Floor, Chicago, IL 60611, USA (e-mail: k-waite@northwestern.edu).*

REFERENCES

1. **Kalichman SC, Ramachandran B, Catz S.** Adherence to combination antiretroviral therapies in HIV patients of low literacy. J Gen Intern Med. 1999;14(5):267–73.
2. **Wolf MS, Davis TC, Osborn CY, Skipkauskas S, Bennett CL, Makoul G.** Literacy, self-efficacy, and HIV medication adherence. Patient Educ Couns. 2007;65(2):253–60.
3. **Miller LG, Liu H, Hays RD, et al.** Knowledge of antiretroviral regimen dosing and adherence: a longitudinal study. Clin Infect Dis. 2003;36: 514–18.
4. **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. 2006;21(8):835–40.
5. **Pignone MP, Dewalt D.** Literacy and health outcomes Is adherence the missing link? J Gen Intern Med. 2006;21(8):896–7.
6. **Gazmararian JA, Kripalani S, Miller MJ, Echt KV, Ren K, Rask K.** Factors associated with medication refill adherence in cardiovascular-related diseases: a focus on health literacy. J Gen Intern Med. 2006;21 (12):1215–21.
7. **Paasche-Orlow MK, Wolf MS.** the causal pathways linking health literacy to health outcomes. Am J Health Behav. 2007;31(Suppl 1): S19–26.
8. **Crandall CS, Coleman R.** AIDS-related stigmatization and the disruption of social relationships. Pers Relatsh. 1992;9:163–77.
9. **Moneyham L, Seals B, Demi A, Sowell R, Cohen L, Guillory J.** Perceptions of stigma in women infected with HIV. AIDS Patient Care STDS. 1996;10:162–7.
10. **Herek GM, Capitanio JP.** Aids and sexual prejudice. Am Behav Sci. 1999;42:1130–47.
11. **Herek GM, Capitanio JP, Widaman KF.** HIV-related stigma and knowledge in the United States: Prevalence and trends. Am J Public Health. 2002;92:371–7.
12. **Blendon RJ, Donelan K.** Discrimination against people with AIDS: The public's perspective. New Engl J Med. 1988;319:1022–6.
13. **Burris S.** Studying the legal management of HIV-related stigma. Am Behav Sci. 1999;42:1229–43.
14. **Craft SM, Serovich JM.** Family-of-origin factors and partner violence in the intimate relationships of gay men who are HIV positive. J Interpers Violence. 2005;20(7):777–91.
15. **D'Augelli AR.** AIDS fears and homophobia among rural nursing personnel. AIDS Educ Prev. 1989;1:277–84.
16. **Derlega VJ, Lovejoy D, Winstead BA.** Personal accounts on disclosing and concealing HIV-positive test results: weighing the benefits and risks. In Derlega VJ, Barbee AP, eds. HIV and Social Interaction. Thousand Oaks, CA: Sage: 1998:147–64.
17. **Goldin CS.** Stigmatization and AIDS: critical issues in public health. Soc Sci Med. 1994;39:1359–66.
18. **Heckman TG, Somlai AM, Kalichman SC, Franzoi SL, Kelly JA.** Psychosocial differences between urban and rural people living with HIV/AIDS. J Rural Health. 1998;14:138–45.
19. **Liebschutz JM, Geier JL, Horton NJ, Chuang CH, Samet JH.** Physical and sexual violence and health care utilization in HIV-infected persons with alcohol problems. AIDS Care. 2005;17:566–78.
20. **Marzuk PM, Tierney H, Tardiff K, et al.** Increased risk of suicide in persons with AIDS. JAMA. 1988;259:1333–7.
21. **Chesney MA, Smith AW.** Critical delays in HIV testing and care: The potential role of stigma. Am Behav Sci. 1999;42:1162–74.
22. **Golin C, Isasi F, Bontempi JB, Eng E.** Secret pills: HIV-positive patients' experiences taking antiretroviral therapy in North Carolina. AIDS Educ Prev. 2002;14(4):318–29.
23. **Rintamaki LS, Davis TC, Skipkauskas S, Bennett CL, Wolf MS.** Social stigma and HIV medication adherence. AIDS Patient Care STDS. 2006;20(5):359–68.
24. **Fogarty L, Roter D, Larson S, Burke J, Gillespie J, Levy R.** Patient adherence to HIV medication regimens: a review of published and abstract reports. Patient Educ Couns. 2002;46(2):93–108.
25. **Singh N, Berman SM, Swindells S, et al.** Adherence of immunodeficiency virus infected patients to antiretroviral therapy. Clin Infect Dis. 1999;29(4):824–30.
26. **Parikh NS, Parker R, Nurss JR, Baker DW, Williams MV.** Shame and health literacy: the unspoken connection. Patient Educ Couns. 1996;27:33–9.
27. **Wolf MS, Williams MV, Parker RM, Parikh NS, Nowlan AW, Baker DW.** Patients' shame and attitudes toward discussing the results of literacy screening. J Health Comm. 2007;12:1–12.
28. **Stirratt MJ, Remien RH, Smith A, et al.** The role of HIV serostatus disclosure in antiretroviral medication adherence. AIDS Behav. 2006;10 (5):483–93.
29. **Peretti-Watel P, Pierret J, Lert F, Obadia Y.** The Vespa Group Management of HIV-related stigma and adherence to HAART: Evidence from a large representative sample of outpatients attending French hospitals. AIDS Care. 2006;18(3):254–61.
30. **Kalichman SC, Rompa D.** Functional health literacy is associated with health status and health-related knowledge in people living with HIV-AIDS. J Acquir Immune Defic Syndr Hum Retrovirol. 2000;25(4):337–44.
31. **Osborn CY, Paasche-Orlow MK, Davis TC, Wolf MS.** Health literacy: An overlooked factor in understanding HIV health disparities. Am J Prev Med. 2007;33(5):374–8.
32. **Wolf MS, Davis TC, Arozullah A, et al.** Relation between literacy and HIV treatment knowledge among patients on HAART regimens. AIDS Care. 2005;17(7):863–73.
33. **Demasi RA, Graham NM, Tolson JM, et al.** Correlation between self-reported adherence to highly active antiretroviral therapy (HAART) and virologic outcome. Adv Ther. 2001;18(4):163–73.
34. **Demasi RA, et al.,** Self-reported adherence to HAART and correlation with HIV RNA: initial results with the patient medication adherence questionnaire. Paper presented at the sixth conference on retroviruses and opportunistic infections (1999).
35. **Frable DE, Blackstone T, Scherbaum C.** Marginal and mindful: Deviants in social interactions. J Pers Soc Psychol. 1990;59:140–9.

36. **Limandri BJ.** Disclosure of stigmatizing condition: The discloser's perspective. Arch of Psychiatr Nurs. 1989;3:69–78.
37. **Goffman E.** Stigma: Notes on the Management of Spoiled Identity. Englewood Cliffs, NJ: Prentice-Hall; 1963.
38. **Hiller DV.** Overweight as master status: A replication. J Psychol. 1982;110:107–13.
39. **Davis TC, Michielutte R, Askov EN, Williams MV, Weiss BD.** Practical assessment of adult literacy in healthcare. Health Educ Behav. 1998;25:613–24.
40. **Davis TC, et al.** Literacy testing in health care research in understanding health literacy: implications for medicine, public health. In: Schwartzberg JG, VanGeest JB, Wang CC, eds. AMA Press: Chicago, IL; 2004:157–79.
41. **Parker RM, Baker DW, Williams MV, Nurss JR.** The test of functional health literacy in adults: a new instrument for measuring patients' literacy skills. J Gen Intern Med. 1995;10:537–41.
42. **Baron RM, Kenny DA.** The moderator-mediator distinction in social psychology research: conceptual, strategic, and statistical considerations. J Pers Soc Psychol. 1986;51:1173–82.
43. **Vanable OA, Carey MP, Blair DC, Littlewood RA.** Impact of HIV-related stigma on health behaviors and psychological adjustment among HIV-positive men and women. AIDS Behav. 2006;10(5):473–82.
44. **Berger BE, Ferran CE, Lashley FR.** Measuring stigma in people with HIV: psychometric assessment of the HIV stigma scale. Res Nurs Health. 2001;24(6):518–29.
45. **Glass TR, De Geest S, Weber R, et al.** Correlates of self-reported nonadherence to antiretroviral therapy in HIV infected patients: the Swiss HIV Cohort Study. J Acquir Immune Defic Syndr. 2006;41:385–92.
46. **Gallant JE, DeJesus E, Arribas JR, et al.** Tenofovir DF, emtricitabine, and efavirenz vs. zidovudine, lamivudine, and efavirens for HIV. N Engl J Med. 2006;354:251–60.
47. **Doak CC, Doak LG, Root JH.** Teaching Patients with Low Literacy Skills. 2nd ed. Philadelphia: Lippincott; 1996.
48. **Gerber BS, Brodsky IG, Lawless KA, et al.** Implementation and evaluation of a low-literacy diabetes education computer multimedia application. Diabetes Care. 2005;28:1574–80.
49. **Davis TC, Fredrickson DD, Arnold C, Murphy PW, Herbst M, Bocchini JA.** A polio immunization pamphlet with increased appeal and simplified language does not improve comprehension to an acceptable level. Patient Educ Couns. 1998;35:25–3.
50. **Davis TC, Williams MV, Marin E, Parker RM, Glass J.** Health literacy and cancer communication. CA Cancer J Clin. 2002;52:134–49.
51. **Lorig KR, Holman H.** Self-management education: history, definition, outcomes, and mechanisms. Ann Behav Med. 2003;26:1–7.
52. **Gifford AL, Laurent DD, Gonzales VM, Chesney MA, Lorig KR.** Pilot randomized trial of education to improve self-management skills of men with symptomatic HIV/AIDS. J Acquir Immune Defic Syndr Hum Retrovirol. 1998;18:136–44.
53. **Bodenheimer T, Lorig K, Holman H, Grumbach K.** Patient self-management of chronic disease in primary care. J Amer Med Assoc. 2002;288:2469–75.
54. **Heckman TG, Barcikowski R, Ogles B, et al.** A telephone-delivered coping improvement group intervention for middle-aged and older adults living with HIV/AIDS. Ann Behav Med. 2006;32(1):27–38.
55. **Jones DL, McPherson-Baker S, Lydston D, et al.** Efficacy of a group medication adherence intervention among HIV positive women: the SMART/EST Women's Project. AIDS Behav. 2007;11(1):79–86.
56. **Makoul G.** The Segue framework for teaching and assessing communication skills. Pat Ed Counsel. 2001;45:23–34.
57. **Murray MD, Young J, Hoke S, et al.** Pharmacist intervention to improve medication adherence in heart failure: a randomized trail. Ann Intern Med. 2007;146(10):714–25, May.

Office-Based Management of Opioid Dependence with Buprenorphine: Clinical Practices and Barriers

Alexander Y. Walley, MD, MSc[1,2], Julie K. Alperen, DrPH[3], Debbie M. Cheng, ScD[1,4], Michael Botticelli[2], Carolyn Castro-Donlan[2], Jeffrey H. Samet, MD, MA, MPH[1,5], and Daniel P. Alford, MD, MPH[1]

[1]Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Boston Medical Center and Boston University School of Medicine, Boston, MA, USA; [2]Bureau of Substance Abuse Services, Massachusetts Department of Public Health, Boston, MA, USA; [3]Department of Epidemiology, Harvard School of Public Health, Boston, MA, USA; [4]Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA; [5]Department of Social and Behavioral Sciences, Boston University School of Public Health, Boston, MA, USA.

BACKGROUND: Buprenorphine is a safe, effective and underutilized treatment for opioid dependence that requires special credentialing, known as a waiver, to prescribe in the United States.

OBJECTIVE: To describe buprenorphine clinical practices and barriers among office-based physicians.

DESIGN: Cross-sectional survey.

PARTICIPANTS: Two hundred thirty-five office-based physicians waivered to prescribe buprenorphine in Massachusetts.

MEASUREMENTS: Questionnaires mailed to all waivered physicians in Massachusetts in October and November 2005 included questions on medical specialty, practice setting, clinical practices, and barriers to prescribing. Logistic regression analyses were used to identify factors associated with prescribing.

RESULTS: Prescribers were 66% of respondents and prescribed to a median of ten patients. Clinical practices included mandatory counseling (79%), drug screening (82%), observed induction (57%), linkage to methadone maintenance (40%), and storing buprenorphine notes separate from other medical records (33%). Most non-prescribers (54%) reported they would prescribe if barriers were reduced. Being a primary care physician compared to a psychiatrist (AOR: 3.02; 95% CI: 1.48–6.18) and solo practice only compared to group practice (AOR: 3.01; 95% CI: 1.23–7.35) were associated with prescribing, while reporting low patient demand (AOR: 0.043, 95% CI: 0.009–0.21) and insufficient institutional support (AOR: 0.37; 95% CI: 0.15–0.89) were associated with not prescribing.

CONCLUSIONS: Capacity for increased buprenorphine prescribing exists among physicians who have already obtained a waiver to prescribe. Increased efforts to link waivered physicians with opioid-dependent patients and initiatives to improve institutional support may mitigate barriers to buprenorphine treatment. Several guideline-driven practices have been widely adopted, such as adjunctive counseling and monitoring patients with drug screening.

Published online July 1, 2008

KEY WORDS: opioid dependence; buprenorphine; medication assisted treatment.
J Gen Intern Med 23(9):1393–8
DOI: 10.1007/s11606-008-0686-x

INTRODUCTION

Although opioid dependence is steadily increasing in the United States, the number of federally licensed methadone maintenance treatment slots is unchanged at approximately 250,000 and unevenly distributed geographically.[1] Buprenorphine is a safe, effective medication for opioid dependence that is associated with increased treatment retention, reduced illicit opioid use, reduced opioid craving, increased survival, and few adverse effects in research and community office-based settings.[2–12] With the enactment of the Drug Addiction Treatment Act (DATA) of 2000 and the Food and Drug Administration's approval of sublingual buprenorphine for the treatment of opioid dependence in October 2002, office-based physicians in the United States, such as primary care physicians and psychiatrists, gained the opportunity to treat opioid-dependent patients with buprenorphine, commonly referred to as office-based opioid treatment (OBOT).[13]

To provide OBOT with buprenorphine, DATA 2000 requires physicians to obtain a waiver from the federal Center for Substance Abuse Treatment (CSAT). Prescribing physicians are responsible for providing patients with or referring patients to substance abuse counseling as well as developing linkages to other addiction treatment programs (e.g., methadone maintenance treatment). CSAT-issued guidelines form the core curriculum of treatment practices for the required 8-hour certification course.[14] Sublingual buprenorphine is indicated for medically supervised taper (detoxification) or maintenance treatment in opioid-dependent patients. The guidelines recommend that dosing during the 1st day of induction be observed in a medical setting, such as the office. The guidelines recommend monitoring treatment adherence with drug toxicology screening, and pill counts. To decrease abuse and diversion, the use of the sublingual tablet coformulation of buprenorphine/naloxone ("combo" tablet) is recommended in all cases except for pregnant patients and during the first part of supervised induction for

patients who are dependent on long-acting opioids. In these cases, the sublingual tablet of buprenorphine alone ("mono" tablet) is indicated. Consistent with federal public health regulations that require specific protections of drug and alcohol treatment information,[15] procedures should be established to ensure the privacy and confidentiality of patients treated with buprenorphine, such as storing OBOT records separately from other general medical information.

With buprenorphine as the first opioid agonist medication FDA-approved for OBOT, generalist physicians are taking on a new role in addiction treatment that requires additional guideline-driven training and certification. Because the system to certify and waiver physicians to prescribe buprenorphine is unique, it is important to examine the barriers waivered physicians encounter in prescribing this medication and how practices compare to guidelines. Previous examinations of barriers to providing OBOT with buprenorphine were conducted within a year of its approval and were focused on addiction specialists.[16,17] In September 2003, Kissin et al. surveyed a random national sample of 545 addiction specialists who had obtained a waiver, finding that 58% of those waivered, prescribed. Significant predictors of prescribing included longer time since obtaining waiver, working in a solo or "individual" practice, and working in a specialty substance abuse treatment program. Predictors of not prescribing included being a psychiatrist.

Because of the substantial mortality and morbidity from opioid dependence,[18] the Massachusetts Department of Public Health (MA DPH) surveyed all 356 waivered physicians in October and November of 2005 about their practices and barriers to office-based opioid treatment with buprenorphine. This study advances current understanding because we surveyed both the barriers to prescribing buprenorphine and treatment practices among substantial numbers of physicians without addiction specialty certification who were waivered to prescribe buprenorphine 3 years after it was initially available. We report the findings from this survey.

METHODS

Population

In October of 2005, MA DPH mailed two questionnaires, one for prescribers and one for non-prescribers, to all 356 physicians in Massachusetts waivered by CSAT to prescribe buprenorphine. Non-respondents were sent a second mailing in November 2005. Mailings included a cover letter from MA DPH explaining that only one of the two mailed questionnaires should be completed based on whether or not the physician was currently prescribing buprenorphine. A self-addressed, stamped envelope was included in the mailing. No personal demographic information was collected and no compensation was offered for completing the questionnaires. Respondents who indicated on their questionnaire that they did not work in an office-based setting were excluded from analysis, because of our focus on describing OBOT.

Data Collection

Both questionnaires included questions about medical specialty, practice setting, addiction society certification, and barriers to prescribing. Respondents were asked about nine barriers to prescribing (for prescribers, "Have you experienced any of the following barriers to the provision of buprenorphine treatment?" and for non-prescribers, "Why do you not currently prescribe buprenorphine?"), including stigma among office staff, "payment issues," "pharmacy issues," low patient demand for treatment, insufficient physician or staff knowledge, and lack of nursing, office, and institutional support. In addition, a space to specify "other" barriers was provided. The prescriber questionnaire inquired about the following treatment practices: the number of patients currently treated; prescribing indication (detox or maintenance); buprenophine induction location (office versus home); the use of substance abuse counseling; the availability of methadone maintenance programs for referral of the use of monitoring practices, such as drug screens or pill counts; indication for the use of the mono tablet; and whether OBOT patient information is stored separately from other patient medical records. The non-prescriber questionnaire asked about past prescribing and plans to prescribe in the future.

Statistical Analysis

The dependent variable of interest was prescriber status (prescriber vs. non-prescriber). We performed frequencies and proportions for each of the questionnaire responses above, overall and by prescriber status. We dichotomized the medical specialty variable into psychiatrists and primary care physicians (general internists, family practitioners, or pediatricians), as these disciplines accounted for nearly all waivered physicians. We also dichotomized the practice setting variable (i.e., group vs. solo), collapsing those who reported both group and solo practice into the group practice category. We conducted bivariate comparisons of these variables and the barrier variables for prescribers and non-prescribers using chi-square tests. To identify factors associated with prescribing buprenorphine, we fit multiple logistic regression models, including the following independent variables: medical specialty, practice setting, addiction society certification, and seven potential barriers to prescribing (insufficient nursing/office support, lack of institutional support, low demand, insufficient physician knowledge, insufficient staff knowledge, payment issues, and pharmacy issues). Correlation analyses were performed prior to performing regression analyses to identify potential variables that may be collinear ($r>0.4$). Because of high correlation between the nursing support barrier and the office support barrier ($r=0.56$), we combined these into a single variable (insufficient nursing or office support) for the logistic regression. We did not include the office staff stigma barrier in the regression, because no non-prescribers reported this barrier. This study was approved by the institutional review board at Boston Medical Center. All statistical analyses were done using SAS 9.1.

RESULTS

Of the 356 waivered physicians who were sent a survey, 235 (66%) were included in our analysis. Of the 121 physicians who were excluded from analyses, 20 questionnaires were returned because of wrong address, 80 questionnaires were non-respondents, and 21 respondents were excluded because they did not practice in an office setting (Fig. 1). The survey response rate, excluding wrong addresses and non-office based physicians, was 75% (235/315).

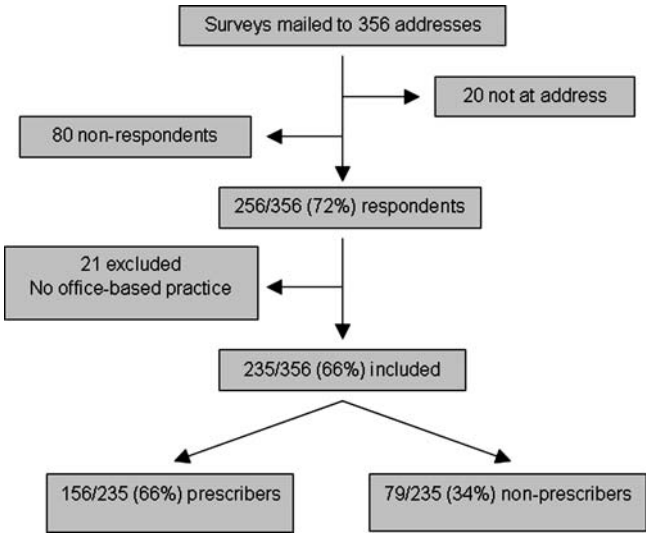


Figure 1. Survey response.

Characteristics and Barriers Among Prescribers and Non-prescribers

Among the 235 physicians in the study sample, 156/235 (66%) were prescribers and 79/235 (34%) were non-prescribers. The characteristics and barriers reported by the total sample and by prescriber status are presented in Table 1. Respondents were from psychiatry (54%) and primary care specialties (44%), in some group practice setting (74%), or only solo practice (26%). Addiction society certification was held by 24%. The more common barriers among prescribers were payment issues, insufficient nursing support, insufficient office support, lack of institutional support, and pharmacy issues. The more common barriers for non-prescribers were insufficient office support, insufficient nursing support, lack of institutional support, insufficient staff knowledge, and low demand. About half of prescribers reported at least one barrier, whereas two-thirds of non-prescribers reported at least one barrier.

OBOT Practices

Results of the prescriber survey are presented in Table 2. Prescribers reported currently treating a median of 10 (interquartile range 3–25.5) patients and a mean of 14 patients. Only 8% of prescribers provided detoxification treatment alone. Forty-three percent of prescribers have at least some of their patients on complete buprenorphine induction at home (i.e., not directly observed). Substance abuse counseling was mandated by 79%, with 66% offering individual and 39% offering group counseling in their practice, and 57% referring to counseling elsewhere. To address opioid-dependent patients who failed or did not qualify for buprenorphine, 86% of respondents stated that they had the capacity to refer patients to methadone maintenance treatment and 40% had made such referrals. Monitoring practices for illicit drug use and appropriate adherence to prescribed medications included: pill counts (43%) and drug screens (82%), with over half reporting "observed" collection of drug screens. The "mono" tablet not co-formulated with naloxone was used by 29% for the following indications: 10% induction, 13% in pregnant patients, and 11% for "patient preference." Only one-third reported maintaining OBOT patient information separately from other medical information. Four-fifths of prescribers accepted insurance for OBOT.

Non-prescribers and OBOT

Of those non-prescribers who had never prescribed buprenorphine, 54% (33/61) reported they will prescribe in the future if the barriers are diminished. Reasons that respondents gave for not prescribing beyond barriers specifically queried included the following: the induction period was "too demanding," the 30-patient limit, insufficient time or space in current practice, and lack of another physician for backup. Of the non-prescribers who had prescribed in the past, 67% (10/15) reported that they would prescribe in the future if barriers were diminished.

Factors associated with prescriber status

Results of logistic regression analyses modeling prescriber status are presented in Table 3. Being a primary care physician (OR 3.02 95% CI: 1.48–6.18) and being in a solo practice only (OR 3.01 95% CI: 1.23–7.35) were factors that significantly increased the odds of prescribing in adjusted analyses. Barriers significantly associated with decreased odds of prescribing in adjusted analyses included insufficient institutional support (OR 0.37 95% CI: 0.15–0.89) and low patient demand for buprenorphine (OR 0.04 95% CI: 0.01–0.21).

Table 1. Characteristics of 235 Office-Based Respondents Waivered to Prescribe Buprenorphine in Massachusetts, Overall and by Prescriber Status

	Total n (%)	Prescribers n (%)	Non-prescribers n (%)
Total	235 (100)	156 (66)	79 (34)
Specialty*			
Psychiatrist	126 (54)	74 (47)	52 (67)
Primary care†	102 (44)	78 (50)	24 (31)
Internist	61 (26)	45 (29)	16 (21)
Family medicine	32 (14)	27 (17)	5 (6)
Pediatrician	9 (4)	6 (4)	3 (4)
Other‡	6 (3)	4(3)	2 (3)
Practice setting§			
Solo	58 (26)	46 (30)	12 (16)
Group	169 (74)	107 (70)	62 (84)
Addiction society certified‖	55 (24)	41 (27)	14 (18)
Barriers¶			
Insufficient nursing support	46 (20)	25 (16)	21 (27)
Insufficient office support	44 (19)	20 (13)	24 (30)
Lack of institutional support	38 (16)	19 (12)	19 (24)
Office staff stigma	11 (5)	11 (7)	0 (0)
Low demand	17 (7)	3 (2)	14 (18)
Insufficient physician knowledge	7 (3)	5 (3)	2 (3)
Insufficient staff knowledge	28 (12)	11 (7)	17 (22)
Payment issues	40 (17)	32 (21)	8 (10)
Pharmacy issues	19 (8)	18 (12)	1 (1)
One or more barriers	130 (55)	77 (49)	53 (67)

* N=234, 1 with missing data
†Primary care is the sum of internists, family medicine, and pediatricians. "Other" excluded
‡Other specialties included "anesthesia/pain medicine" × 2, "retired surgeon,"ambulatory/urgent care medicine," "outpatient addiction/mental health non-psychiatrist," "emergency medicine"
§*N=227, 8 with missing data*
‖*N=232, 3 with missing data*
¶*N=235, none missing*

Table 2. Treatment Practices of 156 Office-Based Buprenorphine Prescribers in Massachusetts

Treatment practices	N=156
Number of patients treated, mean	13.7
Median (interquartile range)	10 (3–25.5)
Detox or maintenance* n (%)	
Detox only	12 (8)
Maintenance only	62 (41)
Detox and maintenance	77 (51)
Induction site†	
Office only (observed)	71 (47)
Home only (unobserved)	31 (21)
Office and home	32 (21)
Inpatient only	15 (10)
Substance abuse counseling	
Mandatory counseling	120 (79)
Individual counseling offered in practice	100 (66)
Group counseling offered in practice	59 (39)
Offer referral to counseling	87 (57)
Methadone program available for referral	131 (86)
Made referrals to methadone program	61 (40)
Monitoring practices:	
Pill counts	67 (43)
Drug screens	128 (82)
Observed drug screens	68 (44)
Unobserved drug screens	60 (38)
Used mono tablet (buprenorphine alone) for:	
Induction	15 (10)
Pregnant patients	21 (13)
Patient preference	17 (11)
OBOT notes stored separate from other records	46 (33)
Accept insurance for buprenorphine	119 (80)

*N=151, 5 with missing data
†N=149, 7 with missing data

DISCUSSION

This survey of waivered physicians in Massachusetts revealed several important findings about who is prescribing buprenorphine and how they are doing it. Given the median number of active patients among prescribers (i.e., ten), substantial treatment capacity among current prescribers remains. Thus, efforts to increase OBOT treatment could be directed to both waivered physicians who already prescribe but have further capacity by regulations as well as those who do not prescribe. The lack of office and nursing support noted as common barriers by both prescribers and non-prescribers is evidence that for many providers adding OBOT with buprenorphine to one's practice requires increased administrative and clinical resources. An example of a successful collaborative care model was recently described.[3] As prescribers commonly identified payment and pharmacy issues as barriers, it is likely that increasing insurance coverage for buprenorphine and making it more available in pharmacies would help prescribers treat more patients. Increasing prescribing among non-prescribers will likely require improved top-down institutional support and improved systems that match patients seeking treatment to waivered physicians.

Determining why psychiatrists were less likely to prescribe than physicians in primary care specialties warrants further investigation. This reluctance of psychiatrists to prescribe buprenorpine was noted previously in a national survey of 1,203 psychiatrists conducted before buprenorphine was released for OBOT, where four-fifths of all respondents, including 43% of those certified in addiction psychiatry, reported they would not be comfortable providing OBOT.[19] The increased likelihood of prescribing we found among primary care physicians compared to psychiatrists may be evidence that the DATA 2000 legislation has encouraged office-based treatment beyond specialty practices and into primary care. According to the CSAT guidelines, DATA 2000 "promises to bring opioid addiction care into the mainstream of medical practice."[14]

We found that being in solo practice versus being in a group practice was also associated with prescribing buprenorphine. Wolinsky and Friedson have described a trade off between greater resources and greater autonomy for physicians who choose to work in group or solo practices, respectively.[20] We expected that group practices could provide more administrative support to waivered physicians who would therefore be more likely to prescribe. However, it is likely that while group practices are better resourced, they present more bureaucratic or administrative barriers to instituting a new treatment such as OBOT. Furthermore, early regulations restricted not only each individual physician to 30 patients, but each group practice to 30 total patients, which likely reduced the incentive for group practices to support providing buprenorphine treatment over competing priorities.

Clinical practices of prescribers were largely consistent with the substance and spirit of buprenorphine training and the CSAT guidelines in that they conform to the 30-patient limit that was in effect at the time of this survey, substance abuse counseling was available and being offered, and monitoring of adherence and relapse through drug screens and pill counts occured widely among prescribers.

Although these OBOT practices were CSAT guideline-driven, others were not. Substantial numbers of physicians use unobserved home induction where patients start buprenorphine at home, usually with telephone support from a nurse or physician. A successful home induction protocol has been described,[3] but is

Table 3. Unadjusted and Adjusted Logistic Regression Models for Prescribing Buprenorphine Among Waivered Physicians in Massachusetts*

	Unadjusted odds ratio	95% CI	P-value	Adjusted odds ratio	95% CI	P-value
Primary care vs. psychiatry	**2.28**	**1.28–4.07**	0.005	**3.02**	**1.48–6.18**	0.002
Solo vs. group practice	**2.20**	**1.08–4.49**	0.027	**3.01**	**1.23–7.35**	0.016
Addiction society certified vs. not certified	1.66	0.84–3.27	0.145	1.57	0.68–3.61	0.286
Barriers						
Insufficient nursing or office support	**0.41**	**0.22–0.74**	0.003	0.76	0.33–1.74	0.518
Lack of institutional support	**0.43**	**0.22–0.89**	0.022	**0.37**	**0.15–0.89**	0.026
Low demand	**0.09**	**0.03–0.33**	<0.001	**0.04**	**0.01–0.21**	<0.001
Insufficient physician knowledge	1.27	0.24–6.72	0.775	2.95	0.25–34.9	0.390
Insufficient staff knowledge	**0.28**	**0.12–0.62**	0.002	0.45	0.15–1.37	0.160
Payment issues	**2.29**	**1.00–5.24**	0.050	2.00	0.73–5.48	0.177
Pharmacy issues	**10.17**	**1.33–77.7**	0.025	7.76	0.88–68.3	0.065

*N=219 for this model. Complete data not available for 16 respondents. Adjusted odds ratios from model that includes all variables in the table

not part of the CSAT guidelines. Some physicians prescribe the mono tablet for "patient preference." This is not an appropriate indication because the mono tablet is more likely to be abused by crushing it and injecting it. Thus, it is more likely to be diverted and has a higher risk of contributing to overdoses.[21] Only one third of prescribers store their notes separately from other medical information, which is a practice not specifically required, but may facilitate compliance with federal confidentiality requirements.

Because this study targeted all physicians eligible to provide OBOT with buprenorphine in a single state almost 3 years after buprenorphine was available, it adds to and supports previous examinations of treatment practices and barriers.[22] We found that two thirds of waivered physicians provided OBOT with buprenorphine, confirming preliminary national estimates.[1,16,17] As in our study, Kissin et al. found that factors associated with not prescribing buprenorphine included being a psychiatrist and working in a setting other than a solo practice. Common barriers noted in this study included concern around the induction logistics, availability of the medication, and the 30-patient limit per physician and per practice that was in force at that time.

A survey of 375 physicians attending HIV educational conferences in 2006 found 25% had obtained a waiver to prescribe buprenorphine, but only 6% had ever prescribed.[23] As in our study, the provider specialty was significantly associated with likelihood of prescribing buprenorphine. Among the HIV providers, general internists were more likely to prescribe than family medicine or infectious disease physicians. Common barriers to providing care noted by waivered respondents included deficits in knowledge about opioid treatment, lack of immediate telephone access to an addiction expert, inability to refer to a substance abuse treatment program, concern about resistance from staff or colleagues, and fear of taking on increased medicolegal risks, overly complicated patients, and issues of medication diversion. Similar knowledge deficits were not commonly reported in our study, though lack of nursing, office, and institutional support were.

The issue of lack of institutional support as a barrier deserves further study. Our survey did not clearly define lack of institutional support, whether it is a barrier from group practice, insurance carrier, hospital or clinic administration. Our findings do suggest that the impact of lack of institutional support is independent of whether a waivered physician is in a solo or group practice. Potential improvements in institutional support are suggested by a 2003 survey of primary care and HIV clinic directors in New York examining the barriers to providing OBOT with buprenorphine.[24] This study found 60% would be likely to provide OBOT with buprenorphine if training was offered. Clinic characteristics associated with increased likelihood of prescribing included providing HIV specialty care, having a secure site to store narcotics, having immediate telephone access to an addiction expert, and receiving continuing medical education credits for training.

Our study has some limitations. First, the number of patients physicians are permitted to treat has evolved, and thus our results may not fully reflect current conditions. We conducted the survey 1 month after the 30-patient limit on each group practice was lifted. Thus, some group practices may have been unwilling to commit physician time or resources to so few patients at the time of the survey, which may explain why physicians in group practices were less likely to be prescribers. Furthermore, in January of 2007, the 30-patient limit per physician was increased to 100 patients for approved physicians prescribing for greater than 1 year. As this Massachusetts sample shows that most prescribers are not close to the 30-patient limit, the impact of increasing the limit to 100 is unlikely to be immediate. Another limitation is that the survey instrument provided nine barriers for respondents to endorse, yet important barriers may not have been included on the list, such as the 30-patient limit. A third of the non-prescribers reported none of the barriers listed, and thus some barriers to prescribing were not identified in this study.

Addictive disorders, such as opioid dependence, are chronic relapsing brain diseases. Like methadone maintenance, OBOT with buprenorphine is probably most effective as a chronic therapy for opioid dependence.[2] Fully integrating this treatment for a chronic disease into mainstream medical practice is occurring among generalist physicians and will likely be enhanced with substantial systematic, multidisciplinary support. Prescribing practices are largely consistent with guidelines, though more education about home induction and the indications for the mono buprenorphine-only formulation should be incorporated into training. Our study provides evidence that utilizing the exisitng treatment capacity among physicians waivered to provide OBOT awaits the improvement of nursing, office, and institutional support and the resolution of payment and pharmacy issues.

Acknowledgements: *The authors would like to thank Amber Jamanka, Jacqueline Ashba, and Andrew Hanchett for data management and analysis contributions, Colleen Labelle, RN, for assistance with survey design and feedback on the manuscript, and Courtney Pierce for assistance in preparing the manuscript. This project was funded by the Massachusetts Department of Public Health. Parts of this work were presented at the Society of General Internal Medicine New England regional meeting, 23 March 2007, the College on Problems in Drug Dependence annual meeting, 19 June 2007, and the Association for Medical Education and Research in Substance Abuse annual meeting, 9 November 2007.*

Conflict of Interest: *None disclosed.*

Corresponding Author: *Alexander Y. Walley, MD, MSc; Section of General Internal Medicine, Boston Medical Center, 801 Massachusetts Avenue # 2156, Boston, MA 02118, USA (e-mail: awalley@bu.edu).*

REFERENCE

1. Center for Substance Abuse Treatment, Substance Abuse Treatment and Mental Health Services Administration, US Department of Health and Human Services. The determinations report: A report on the physician waiver program established by the Drug Addiction Treatment Act of 2000 ("DATA"). Available at: buprenorphine.samhsa.gov/SAMHSA_Determinations_Report.pdf. Accessed: May 16, 2008.
2. **Kakko J, Svanborg KD, Kreek MJ, Heilig M.** One-year retention and social function after buprenorphine-assisted relapse prevention treatment for heroin dependence in Sweden: a randomised, placebo-controlled trial. Lancet. 2003;361(9358):662–8.
3. **Alford DP, LaBelle CT, Richardson JM, et al.** Treating homeless opioid dependent patients with buprenorphine in an office-based setting. J Gen Intern Med. 2007;22(2):171–6.
4. **Stein MD, Cioe P, Friedmann PD.** Buprenorphine retention in primary care. J Gen Intern Med. 2005;20(11):1038–41.
5. **Mintzer IL, Eisenberg M, Terra M, MacVane C, Himmelstein DU, Woolhandler S.** Treating opioid addiction with buprenorphine-naloxone in community-based primary care settings. Ann Fam Med. 2007;5(2):146–50.
6. **Magura S, Lee SJ, Salsitz EA, et al.** Outcomes of buprenorphine maintenance in office-based practice. J Addict Dis. 2007;26(2):13–23.

7. **Mattick RP, Kimber J, Breen C, Davoli M.** Buprenorphine maintenance versus placebo or methadone maintenance for opioid dependence. Cochrane Database Syst Rev. 2004;(3):CD002207
8. **Johnson RE, Jaffe JH, Fudala PJ.** A controlled trial of buprenorphine treatment for opioid dependence. JAMA. 1992;267(20):2750–5.
9. **Johnson RE, Chutuape MA, Strain EC, Walsh SL, Stitzer ML, Bigelow GE.** A comparison of levomethadyl acetate, buprenorphine, and methadone for opioid dependence. N Engl J Med. 2000;343(18):1290–7.
10. **Fudala PJ, Bridge TP, Herbert S, et al.** Office-Based Treatment of Opiate Addiction with a Sublingual-Tablet Formulation of Buprenorphine and Naloxone. N Engl J Med. 2003;349(10):949–8.
11. **O'Connor PG, Carroll KM, Shi JM, Schottenfeld RS, Kosten TR, Rounsaville BJ.** Three Methods of Opioid Detoxification in a Primary Care Setting: A Randomized Trial. Ann Intern Med. 1997;127(7):526–30.
12. **Ling W, Charuvastra C, Collins JF, et al**. Buprenorphine maintenance treatment of opiate dependence: a multicenter, randomized clinical trial. Addiction. 1998;93(4):475–486.
13. Drug Addiction Treatment Act of 2000. Public Law 106–310, Page 114 STAT.1101-Page 114 STAT.1227. 8-2-0005.
14. Center for Substance Abuse Treatment. Clinical guidelines for the use of buprenorphine in the treatment of opioid addiction. Treatment Improvement Protocol (TIP) Series 40.DHHS Publication No. (SMA) 04–3939. 2004. Rockville, MD.
15. Confidentiality of Alcohol and Drug Abuse Patient Records. Code of Federal Regulations Title 42, Volume 1, Chapter 1, Part 2. Available at: http://www.2hs.gov/ohrp/documents/19750701.pdf. Accessed: May 16, 2008.
16. **Bivol, S.** National poll of physicians on barriers to widespread buprenorphine use. 2003. Boston, Join Together. 8-29-2007.
17. **Kissin W, McLeod C, Sonnefeld J, Stanton A.** Experiences of a national sample of qualified addiction specialists who have and have not prescribed buprenorphine for opioid dependence. J Addict Dis. 2006;25(4):91–103.
18. Center for Health Information, Statistics, Research, and Evaluation, Bureau of Substance Abuse Services, Massachusetts Department of Public Health. Opioids: Trends and Current Status in Massachusetts. Available at: www.mass.gov/Eeohhs2/docs/dph/injury_surveillance/opioid_trends_2007.doc. Accessed: May 16, 2008.
19. **West JC, Kosten TR, Wilk J, et al**. Challenges in increasing access to buprenorphine treatment for opiate addiction. Am J Addict. 2004;13(Suppl 1):S8–S16.
20. **Wolinsky FD.** Why physicians choose different types of practice settings. Health Serv Res. 1982;17(4):399–419.
21. **Gaulier JM, Marquet P, Lacassie E, Dupuy JL, Lachatre G.** Fatal intoxication following self-administration of a massive dose of buprenorphine. J Forensic Sci. 2000;45(1):226–8.
22. **Becker WC, Fiellin DA.** Provider satisfaction with office-based treatment of opioid dependence: a systematic review. Subst Abus. 2005;26(1):15–22.
23. **Cunningham CO, Kunins HV, Roose RJ, Elam RT, Sohler NL.** Barriers to obtaining waivers to prescribe buprenorphine for opioid addiction treatment among HIV physicians. J Gen Intern Med. 2007;22(9):1325–9.
24. **Turner BJ, Laine C, Lin YT, Lynch K.** Barriers and facilitators to primary care or human immunodeficiency virus clinics providing methadone or buprenorphine for the management of opioid dependence. Arch Intern Med. 2005;165(15):1769–76.

Recent drug use, homelessness and increased short-term mortality in HIV-infected persons with alcohol problems

Alexander Y. Walley[a], Debbie M. Cheng[a,d], Howard Libman[c], David Nunes[b], C. Robert Horsburgh Jr.[e], Richard Saitz[a,e,f] and Jeffrey H. Samet[a,g]

Objective: To assess the impact of recent heavy alcohol use, heroin/cocaine use, and homelessness on short-term mortality in HIV-infected persons.

Methods: Survival in a longitudinal cohort of 595 HIV-infected persons with alcohol problems was assessed at 6-month intervals in 1996–2005. The time-varying main independent variables were heavy alcohol use (past 30 days), heroin/cocaine use (past 6 months), and homelessness (past 6 months). Date of death was determined using the Social Security Death Index. Outcomes were limited to deaths occurring within 6 months of last assessment to ensure recent assessments of the main independent variables. Cox proportional hazards models were fit to the data.

Results: Death within 6 months of their last assessment occurred in 31 subjects (5.2%). Characteristics at study entry included mean age 41 years, 25% female, 41% African-American, 24% with CD4 cell count < 200 cells/μl; 41% taking antiretroviral therapy, 30% heavy alcohol use, 57% heroin or cocaine use, and 28% homelessness. Heroin or cocaine use [hazard ratio (HR), 2.43; 95% confidence interval (CI), 1.12–5.30)] and homelessness (HR, 2.92; 95% CI, 1.32–6.44), but not heavy alcohol use (HR, 0.57; 95% CI, 0.23–1.44), were associated with increased mortality in analyses adjusted for age, injection drug use ever, CD4 cell count, and current antiretroviral therapy.

Conclusions: Recent heroin or cocaine use and homelessness are associated with increased short-term mortality in HIV-infected patients with alcohol problems. Optimal management of HIV-infected patients requires regular assessments for drug use and homelessness and improved access to drug treatment and housing.

AIDS 2008, **22**:415–420

Keywords: alcohol, cocaine, heroin, HIV, homeless persons, mortality

Introduction

Since the advent of combination active antiretroviral therapy (ART) in 1996, mortality in HIV-infected patients with access to these drugs has decreased substantially [1–7]. However several studies have noted smaller mortality improvements in patients infected with HIV from injection drug use compared with other transmission routes [1,6,8–11]. Recent drug use [12,13], alcohol use [14,15], and homelessness [16–19] are

From the [a]Clinical Addiction Research and Education (CARE) Unit, the [b]Section of Gastroenterology, Department of Medicine, Boston Medical Center and Boston University School of Medicine, the [c]Divisions of General Medicine and Primary Care and Infectious Diseases, Beth Israel Deaconess Medical Center, Harvard Medical School, the [d]Department of Biostatistics, the [e]Department of Epidemiology, the [g]Department of Social and Behavioral Sciences, and the [f]Youth Alcohol Prevention Center, Boston University School of Public Health, Boston, Massachusetts, USA.

Correspondence to Dr. A.Y. Walley, Section of General Internal Medicine, Boston Medical Center, 91 East Concord 200, Boston, MA 02118, USA.

E-mail: awalley@bu.edu

Received: 24 May 2007; revised: 31 October 2007; accepted: 5 November 2007.

common modifiable conditions in HIV-infected patients. The impact of recent homelessness, alcohol use, or illicit drug use on survival among HIV-infected patients is not clear.

Some cohort studies have found no increased mortality from recent homelessness, alcohol, or drug use [20,21]. While two studies showed increased mortality among subjects using drugs at every follow-up assessment compared with those using drugs at no follow-up assessments [22,23], these studies categorized subjects based on their pattern of use over the cohort, rather than using time-varying analyses allowing subjects' status to change during follow-up time. One study of HIV-infected women found an increased risk of non-AIDS-related mortality with recent injection drug use [11]. None of these studies focused on short-term mortality by restricting the outcome to those deaths occurring within 6 months of the last study assessment.

Because substance use and homelessness are dynamic conditions that change over time, focusing on deaths occurring within a shorter period of time, such as 6 months after they are measured ('short-term mortality'), increases the likelihood of measuring a mortality impact from these conditions when they are active. To address the issue of whether substance use and homelessness, as modifiable characteristics, are associated with mortality, we studied the impact of recent heavy alcohol use, recent heroin or cocaine use, and recent homelessness as time-dependent variables on short-term mortality in HIV-infected patients with current or past alcohol problems.

Methods

Population

A survival analysis was performed using data collected from two prospective cohorts, the HIV–Alcohol Longitudinal Cohort (HIV-ALC) and the HIV-Longitudinal Interrelationships of Viruses and Ethanol (HIV-LIVE) cohort, which have been described in prior publications [24,25]. Both cohorts used identical recruitment sites in Boston, Massachusetts. Entry criteria included confirmed HIV positive antibody test, two or more affirmative answers on the CAGE alcohol screening questionnaire [15,26,27] or diagnosis by a study coinvestigator physician of current or past alcohol abuse or dependence, English or Spanish speaking, and at least one contact person available to assist with follow-up. Between July 1997 and July 2001, HIV-ALC enrolled and followed 349 patients. Between September 2001 and November 2005, the HIV-LIVE study enrolled and followed 400 subjects, 154 of whom had also participated in HIV-ALC. The combined HIV-ALC and HIV-LIVE cohort analyzed in this study included 595 subjects.

Data collection and independent variables

Laboratory and interview data were collected by trained research associates every 6 months. The 30-day timeline followback method was used at each interview to measure alcoholic intake [28–30]. Recent heavy alcohol use was defined according to National Institute on Alcohol Abuse and Alcoholism guidelines: >14 drinks/week or >4 drinks on one occasion for men <65 years of age, and >7 drinks/week or >3 drinks on one occasion for all women and men ≥ 65 years of age. Recent heroin or cocaine use was defined as any use over the last 12 months at study entry and over the last 6 months at follow-up assessments via the Composite International Diagnostic Interview – Short Form: Drug Dependence (CIDI-SF DD) [31]. Recent homelessness was defined as having spent a night in the past 6 months in an overnight shelter or on the street.

Potential confounders defined at study entry included age, sex, race/ethnicity, prior injection drug use, prior suicide attempt, CD4 cell count (<200 cells/μl versus ≥ 200 cells/μl), log of HIV viral load, and time of study entry (prior to August 2001 versus after August 2001). Because of the limited number of outcomes, for multivariable analyses, race/ethnicity was dichotomized as black versus nonblack. Time of study entry was considered as a covariate in order to control for potential improvement in HIV care and access between the study entry period for HIV-ALC and HIV-LIVE.

Potential confounders modeled as time-varying covariates included ART status and adherence over 3 days [32], physical and mental health status using the 12-item Short Form Health Survey [33,34], and depressive symptoms using the Center for Epidemiologic Studies Depression (CES-D) scale [35,36]. ART status was examined using the following categories: not on ART, on ART but not adherent, and on ART and adherent. Adherence was defined as 100% adherence over the 3 previous days. Because of the limited number of outcomes, ART status was dichotomized for the multivariable analyses: on ART versus not on ART. Depressive symptoms were dichotomized using a cut-off of ≥ 23 or greater in the CES-D indicating substantial depressive symptoms.

Assessment of date and cause of death

The primary outcome for this study was time from study enrollment to death. Events were restricted to those deaths that occurred within 6 months after the subjects' last study interview ('short-term' mortality) in order to ensure the availability of recent assessments of substance use and homelessness. The observation time for all other subjects was defined as the time from study enrollment until 6 months following their last study visit.

Dates of death were determined by searching the Social Security Death Index from 1997 to 2005. Causes of death were obtained for subjects who died prior to 2004 via the National Death Index Plus and for subjects who died

between 1 January 2004 and 15 November 2005 from the Massachusetts Registry of Vital Records using copies of death certificates. Three of the study physicians (AW, HL, and DN) independently allocated each death into the following categories: HIV related, liver related, overdose, non-HIV/nonliver-related cancer, other infection related, trauma, cardiovascular, and other (deaths for which there was not enough information or death that did not fit into one of the categories above). After independent categorization, the three authors resolved discrepancies through consensus.

Analysis

Time-varying Cox proportional hazards models were used to examine the association of the main independent variables with short-term mortality. Preliminary unadjusted models were fit for each independent variable. It was not possible to fit a multivariable model that included all the main independent variables and potential confounders because of the modest number of deaths occurring within 6 months of the final study visit. A multivariable model was fitted that included the three main independent variables and covariates that have been demonstrated to be important mortality predictors in HIV-infected patients: age, CD4 cell count < 200 cells/μl, prior injection drug use, and ART use [37]. Hepatitis C infection was not included because of substantial correlation ($r = 0.76$) between this infection and past intravenous drug use among the 399 subjects in the HIV-LIVE cohort for whom HCV infection status was available. The remaining covariates were added one at a time to assess whether any one was a confounder. A change in the parameter estimate of $\geq 10\%$ was used to identify confounding. All analyses were performed using SAS 9.1.3 (SAS Institute, Cary, North Carolina, USA).

Results

Of the 595 study subjects, 31 subjects (5.2%) died within 6 months of their last study interview, and 99 subjects (16.6%) died overall. The mean follow-up time was 2.7 years. Characteristics of all 595 subjects at study entry are shown in Table 1. Across the study follow-up, heavy alcohol use was reported in no assessments, at least one assessment, and at all assessments by 52%, 37%, and 11% of subjects, respectively. The respective proportions for heroin/cocaine use were 49%, 35%, and 16%, respectively; for homelessness they were 62%, 26%, and 11%, respectively.

Unadjusted Cox proportional hazards models for each of the main independent variables and potential confounders are also presented in Table 1. Recent heroin or

Table 1. Characteristics and relative hazards of short-term mortality among 595 HIV-infected patients with alcohol problems.

	All subjects at study entry	Crude HR (95% CI)[a]	Adjusted HR (95% CI)[a,b]
Mean age [years (SD)]	41 (7.43)	1.04 (1.00–1.09)	1.06 (1.01–1.13)
Female [No. (%)]	148 (25)	1.38 (0.61–3.10)	
Race/ethnicity [No. (%)]			
African-American	246 (41)	1.05 (0.43–2.55)	
White	202 (34)	Ref	
Other	147 (25)	1.73 (0.68–4.38)	
African-American versus other		0.80 (0.39–1.65)	
Prior injection drug use [No. (%)]	354 (60)	2.81 (1.15–6.87)	1.75 (0.69–4.43)
Prior suicide attempt [No. (%)]	119 (20)	0.54 (0.19–1.55)	
Depressive symptoms* [No. (%)]	291 (49)	1.41 (0.70–2.86)	
Study entry pre-August 2001 [No. (%)]	349 (59)	1.31 (0.52–3.31)	
Mean SF-12 score (SD)			
PCS*	44 (11.2)	0.96 (0.93–1.00)	
MCS*	39 (14.1)	0.98 (0.95–1.00)	
Mean HIV RNA [log copies/ml (SD)]	2.71 (1.98)	1.20 (0.98–1.46)	
CD4 cell count < 200 cells/μl [No. (%)]	136 (24)	2.43 (1.14–5.15)	3.58 (1.59–8.06)
Medication status [No. (%)]			
Not on ART*	240 (40)	1.88 (0.89–3.99)	
On ART, not adherent*	105 (18)	1.02 (0.28–3.62)	
On ART, adherent*	249 (42)	Ref	
Off ART versus on ART*		1.88 (0.93–3.80)	2.92 (1.33–6.40)
Heavy alcohol use in prior 30 days* [No. (%)]	180 (30)	0.61 (0.25–1.49)	0.57 (0.23–1.44)
Heroin/cocaine use[c,*] [No. (%)]	218 (37)	2.39 (1.18–4.85)	2.43 (1.12–5.30)
Homeless in prior 6 months* [No. (%)]	164 (28)	2.85 (1.34–6.07)	2.92 (1.32–6.44)

HR, hazard ratio; CI, confidence interval; SF-12, 12-item Short-Form Health Survey; PCS, physical component summary; MCS, mental component summary; ART, antiretroviral therapy.

[a]For crude and adjusted hazard models, variables were measured at study entry, except for those marked with *, which were analyzed as time-varying variables.

[b]The adjusted Cox proportional hazards model included age, prior injection drug use, CD4 cell count, off ART versus on ART, heavy alcohol use, heroin or cocaine use, and homelessness.

[c]Heroin or cocaine use in the last 12 months for study entry, heroin or cocaine use in the last 6 months for follow-up visits.

cocaine use [hazard ratio (HR), 2.39; 95% confidence interval (CI), 1.18–4.85] and recent homelessness (HR, 2.85; 95% CI, 1.34–6.07) were both associated with increased mortality. Recent heavy alcohol use was not significantly associated with mortality (HR, 0.61; 95% CI, 0.25–1.49).

In the multivariable Cox proportional hazards model controlling for age, prior injection drug use, CD4 cell count, and ART use, recent heroin or cocaine use (HR, 2.43; 95% CI, 1.12–5.30) and recent homelessness (HR, 2.92; 95% CI, 1.32–6.44), but not recent heavy alcohol use (HR, 0.57; 95% CI, 0.23–1.44), were associated with a higher risk of mortality (Table 1). Further adjustment for health status, depressive symptoms, gender, race, prior suicide attempt, and time of study entry did not reveal confounding of the association between the main independent variables and mortality.

Overdose, liver-related conditions, and HIV-related conditions were the three most common categories for the 31 deaths occurring within 6 months of an assessment and for the 99 deaths overall (Table 2).

Discussion

In this longitudinal cohort of HIV-infected patients with current or past alcohol problems, recent heroin or cocaine use and homelessness were associated with an increased risk of short-term mortality. The major strength of this study is the use of regular reassessments of substance use and homelessness states as time-varying covariates. Substance use and homelessness are often conditions that occur together, with presumably similar causes and effects. Yet we were able to disentangle their impacts by including them, along with known important covariates, such as CD4 cell count and prior injection drug use, in multivariable Cox proportional hazards models.

We did not find a similar impact on mortality risk from recent heavy alcohol use. It may be that heavy alcohol use affects only long-term, not short-term, mortality, acting via chronic liver disease. Alternatively, the short-term mortality risk from heavy alcohol use may be mitigated by a 'sick quitter' phenomenon, where sicker subjects are less likely to be drinkers [38,39].

There are several limitations to our study. First, we had a modest number of deaths occurring within 6 months of last interview, limiting our power and ability to control for all potential confounders in the same model. Nonetheless, this definition of the outcome enabled us to focus on the short-term effects of recent conditions. Second, the generalizability of our findings may be limited because we studied HIV-infected subjects with current or past alcohol problems recruited from one urban area. Third, loss to follow-up may have introduced a bias. For example, subjects who missed study visits may have been more likely to be drinking heavily, using drugs, or homeless. However, this bias would likely have biased results towards the null hypothesis.

Many of the deaths in this cohort, such as those caused by overdose, trauma, and infection, were likely acute, preventable, and directly related to recent drug use and homelessness. Integration of effective substance abuse treatment with HIV care has been demonstrated to be effective [40–43]. Successful housing programs with integrated services for homeless people with mental health or substance abuse problems have been described [44–46]. In addition to providing standard medical treatment, optimal management of HIV-infected patients requires regular assessments for drug use and homelessness and improved access to substance abuse treatment and housing.

Table 2. Causes of death.

Cause of death category	Short-term mortality ($n=31$)[a]	Overall mortality ($n=99$)[b]
HIV-related, No. (%)	5 (16)	15 (15)
Liver-related	6 (19)	19 (19)
Overdose	7 (23)	14 (14)
Infection-related	3 (10)	10 (10)
Non-HIV/Nonliver-related cancer	2 (6)	8 (8)
Trauma	2 (6)	4 (4)
Cardiovascular	1 (3)	1 (1)
Other	5 (16)	28 (28)

[a]Deaths occurring within 6 months of subjects' last study interview.
[b]All deaths identified anytime during the study period.

Acknowledgements

The authors appreciate the contributions of the researchers on the project (Carly Bridden, Caitlin McDonnell, and Nella Mupier), the clinical nurse (Colleen Labelle), the data managers (Emily Quinn and Tara Tripp) and software assistance from Christopher Shanahan.

Sponsorship: This study was supported by the National Institute on Alcohol Abuse and Alcoholism (NIAAA): R01-AA13216 (Clinical Impact of HCV and Alcohol in HIV-Infected Persons), R01-AA11785 (Medication Adherence in Alcohol Abusing HIV Patients), R01-AA10870 (Enhanced Linkage of Alcohol Abusers to Primary Care). This research was conducted in part in the General Clinical Research Center at Boston University School of Medicine, National Center for Research Resources (NCRR) M01 RR00533, and the Beth Israel Deaconess Hospital Clinical Research Center, M01 RR01032. Dr. Walley was supported by the National Institute on Drug Abuse (R25-DA13582)

and National Institute of Allergy and Infectious Diseases (T32-AI52074). Dr. Samet was supported by NIAAA (K24 AA015674).

References

1. Keiser O, Taffe P, Zwahlen M, Battegay M, Bernasconi E, Weber R, *et al*. **All cause mortality in the Swiss HIV Cohort Study from 1990 to 2001 in comparison with the Swiss population.** *AIDS* 2004; **18**:1835–1843.
2. Palella FJ Jr, Delaney KM, Moorman AC, Loveless MO, Fuhrer J, Satten GA, *et al*. **Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators.** *N Engl J Med* 1998; **338**:853–860.
3. Detels R, Munoz A, McFarlane G, Kingsley LA, Margolick JB, Giorgi J, *et al*. **Effectiveness of potent antiretroviral therapy on time to AIDS and death in men with known HIV infection duration. Multicenter AIDS Cohort Study Investigators.** *JAMA* 1998; **280**:1497–1503.
4. Hogg RS, Yip B, Kully C, Craib KJ, O'Shaughnessy MV, Schechter MT, *et al*. **Improved survival among HIV-infected patients after initiation of triple-drug antiretroviral regimens.** *CMAJ* 1999; **160**:659–665.
5. Murphy EL, Collier AC, Kalish LA, Assmann SF, Para MF, Flanigan TP, *et al*. **Highly active antiretroviral therapy decreases mortality and morbidity in patients with advanced HIV disease.** *Ann Intern Med* 2001; **135**:17–26.
6. Whitman S, Murphy J, Cohen M, Sherer R. **Marked declines in human immunodeficiency virus-related mortality in Chicago in women, African Americans, Hispanics, young adults, and injection drug users, from 1995 through 1997.** *Arch Intern Med* 2000; **160**:365–369.
7. Vlahov D, Galai N, Safaeian M, Galea S, Kirk GD, Lucas GM, *et al*. **Effectiveness of highly active antiretroviral therapy among injection drug users with late-stage human immunodeficiency virus infection.** *Am J Epidemiol* 2005; **161**:999–1012.
8. Poundstone KE, Chaisson RE, Moore RD. **Differences in HIV disease progression by injection drug use and by sex in the era of highly active antiretroviral therapy.** *AIDS* 2001; **15**:1115–1123.
9. Moore RD, Keruly JC, Chaisson RE. **Differences in HIV disease progression by injecting drug use in HIV-infected persons in care.** *J Acquir Immune Defic Syndr* 2004; **35**:46–51.
10. Perez-Hoyos S, del Amo J, Muga R, del Romero J, Garcia de Olalla P, Guerrero R, *et al*. **Effectiveness of highly active antiretroviral therapy in Spanish cohorts of HIV seroconverters: differences by transmission category.** *AIDS* 2003; **17**:353–359.
11. Smith DK, Gardner LI, Phelps R, Hamburger ME, Carpenter C, Klein RS, *et al*. **Mortality rates and causes of death in a cohort of HIV-infected and uninfected women, 1993–1999.** *J Urban Health* 2003; **80**:676–688.
12. Centers for Disease Control and Prevention. **HIV diagnoses among injection-drug users in states with HIV surveillance: 25 states, 1994–2000.** *MMWR Morb Mortal Wkly Rep* 2003; **52**:634–636.
13. Centers for Disease Control and Prevention. **Trends in HIV/AIDS diagnoses: 33 states, 2001–2004.** *MMWR Morb Mortal Wkly Rep* 2005; **54**:1149–1153.
14. Petry NM. **Alcohol use in HIV patients: what we don't know may hurt us.** *Int J STD AIDS* 1999; **10**:561–570.
15. Samet JH, Phillips SJ, Horton NJ, Traphagen ET, Freedberg KA. **Detecting alcohol problems in HIV-infected patients: use of the CAGE questionnaire.** *AIDS Res Hum Retroviruses* 2004; **20**:151–155.
16. Smereck GA, Hockman EM. **Prevalence of HIV infection and HIV risk behaviors associated with living place: on-the-street homeless drug users as a special target population for public health intervention.** *Am J Drug Alcohol Abuse* 1998; **24**:299–319.
17. Zolopa AR, Hahn JA, Gorter R, Miranda J, Wlodarczyk D, Peterson J, *et al*. **HIV and tuberculosis infection in San Francisco's homeless adults. Prevalence and risk factors in a representative sample.** *JAMA* 1994; **272**:455–461.
18. Torres RA, Mani S, Altholz J, Brickner PW. **Human immunodeficiency virus infection among homeless men in a New York City shelter. Association with *Mycobacterium tuberculosis* infection.** *Arch Intern Med* 1990; **150**:2030–2036.
19. Metsch LR, McCoy CB, McCoy HV, Shultz JM, Lai S, Weatherby NL, *et al*. **HIV-related risk behaviors and seropositivity among homeless drug-abusing women in Miami, Florida.** *J Psychoact Drugs* 1995; **27**:435–446.
20. Kohli R, Lo Y, Howard AA, Buono D, Floris-Moore M, Klein RS, *et al*. **Mortality in an urban cohort of HIV-infected and at-risk drug users in the era of highly active antiretroviral therapy.** *Clin Infect Dis* 2005; **41**:864–872.
21. Riley ED, Bangsberg DR, Guzman D, Perry S, Moss AR. **Antiretroviral therapy, hepatitis C virus, and AIDS mortality among San Francisco's homeless and marginally housed.** *J Acquir Immune Defic Syndr* 2005; **38**:191–195.
22. Lucas GM, Griswold M, Gebo KA, Keruly J, Chaisson RE, Moore RD. **Illicit drug use and HIV-1 disease progression: a longitudinal study in the era of highly active antiretroviral therapy.** *Am J Epidemiol* 2006; **163**:412–420.
23. Kapadia F, Cook JA, Cohen MH, Sohler N, Kovacs A, Greenblatt RM, *et al*. **The relationship between noninjection drug use behaviors on progression to AIDS and death in a cohort of HIV seropositive women in the era of highly active antiretroviral therapy use.** *Addiction* 2005; **100**:990–1002.
24. Samet JH, Horton NJ, Traphagen ET, Lyon SM, Freedberg KA. **Alcohol consumption and HIV disease progression: are they related?** *Alcohol Clin Exp Res* 2003; **27**:862–867.
25. Libman H, Saitz R, Nunes D, Cheng DM, Richardson JM, Vidaver J, *et al*. **Hepatitis C infection is associated with depressive symptoms in HIV-infected adults with alcohol problems.** *Am J Gastroenterol* 2006; **101**:1804–1810.
26. Buchsbaum DG, Buchanan RG, Centor RM, Schnoll SH, Lawton MJ. **Screening for alcohol abuse using CAGE scores and likelihood ratios.** *Ann Intern Med* 1991; **115**:774–777.
27. Mayfield D, McLeod G, Hall P. **The CAGE questionnaire: validation of a new alcoholism screening instrument.** *Am J Psychiatry* 1974; **131**:1121–1123.
28. Sobell L, Sobell M. Timeline followback: a technique for assessing self-reported alcohol consumption. In *Measuring Alcohol Consumption: Psychosocial and Biological Methods*. Edited by Litten RZ, Allen J. New Jersey: Humana Press; 1992. pp. 41–72.
29. Sobell L, Sobell M. *Alcohol timeline followback (TLFB) Users' Manual*. Toronto, Canada: Addiction Research Foundation; 1995.
30. Fals-Stewart W, O'Farrell TJ, Freitas TT, McFarlin SK, Rutigliano P. **The timeline followback reports of psychoactive substance use by drug-abusing patients: psychometric properties.** *J Consult Clin Psychol* 2000; **68**:134–144.
31. Robins LN, Wing J, Wittchen HU, Helzer JE, Babor TF, Burke J, *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**:1069–1077.
32. Chesney MA, Ickovics JR, Chambers DB, Gifford AL, Neidig J, Zwickl B, *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. Delate T, Coons SJ. **The discriminative ability of the 12-item Short Form Health Survey (SF-12) in a sample of persons infected with HIV.** *Clin Ther* 2000; **22**:1112–1120.
34. Viswanathan H, Anderson R, Thomas J 3rd. **Nature and correlates of SF-12 physical and mental quality of life components among low-income HIV adults using an HIV service center.** *Qual Life Res* 2005; **14**:935–944.
35. Kalichman SC, Rompa D, Cage M. **Distinguishing between overlapping somatic symptoms of depression and HIV disease in people living with HIV-AIDS.** *J Nerv Ment Dis* 2000; **188**:662–670.
36. Radloff L. **The CES-D scale: a self-report depression scale for research in the general population.** *Appl Psychol Measure* 1977; **1**:385–401.
37. Egger M, May M, Chene G, Phillips AN, Ledergerber B, Dabis F, *et al*. **Prognosis of HIV-1-infected patients starting highly active antiretroviral therapy: a collaborative analysis of prospective studies.** *Lancet* 2002; **360**:119–129.

38. Green CA, Polen MR. **The health and health behaviors of people who do not drink alcohol.** *Am J Prev Med* 2001; **21**:298–305.
39. Fillmore KM, Golding JM, Graves KL, Kniep S, Leino EV, Romelsjo A, *et al*. **Alcohol consumption and mortality. I. Characteristics of drinking groups.** *Addiction* 1998; **93**:183–203.
40. Lum PJ, Tulsky JP. **The medical management of opioid dependence in HIV primary care settings.** *Curr HIV/AIDS Rep* 2006; **3**:195–204.
41. Basu S, Smith-Rohrberg D, Bruce RD, Altice FL. **Models for integrating buprenorphine therapy into the primary HIV care setting.** *Clin Infect Dis* 2006; **42**:716–721.
42. Sullivan LE, Barry D, Moore BA, Chawarski MC, Tetrault JM, Pantalon MV, *et al*. **A trial of integrated buprenorphine/naloxone and HIV clinical care.** *Clin Infect Dis* 2006; **43 (Suppl 4)**: S184–S190.
43. Rothman J, Rudnick D, Slifer M, Agins B, Heiner K, Birkhead G. **Co-located substance use treatment and HIV prevention and primary care services, New York State, 1990–2002: A model for effective service delivery to a high-risk population.** *J Urban Health* 2007; **84**:226–242.
44. Clark C, Rich AR. **Outcomes of homeless adults with mental illness in a housing program and in case management only.** *Psychiatr Serv* 2003; **54**:78–83.
45. Tsemberis S, Gulcur L, Nakae M. **Housing First, consumer choice, and harm reduction for homeless individuals with a dual diagnosis.** *Am J Public Health* 2004; **94**:651–656.
46. Martinez TE, Burt MR. **Impact of permanent supportive housing on the use of acute care health services by homeless adults.** *Psychiatr Serv* 2006; **57**:992–999.

AIDS Behav (2008) 12:662–669
DOI 10.1007/s10461-007-9243-6

ORIGINAL PAPER

Implications of Cannabis Use and Heavy Alcohol Use on HIV Drug Risk Behaviors in Russian Heroin Users

Alexander Y. Walley · Evgeny M. Krupitsky · Debbie M. Cheng · Anita Raj · Erika M. Edwards · Carly Bridden · Valentina Y. Egorova · Edwin E. Zvartau · George E. Woody · Jeffrey H. Samet

Received: 20 November 2006 / Accepted: 16 April 2007 / Published online: 9 May 2007

Abstract Cannabis and heavy alcohol use potentially increase HIV transmission by increasing risky drug behaviors. We studied 404 subjects entering treatment for heroin dependence, in St. Petersburg, Russia. We used the HIV Risk Assessment Battery (RAB) drug subscale to measure risky drug behavior. Although all heavy alcohol users had risky drug behaviors, their drug RAB scores did not differ from non-heavy alcohol users in unadjusted or adjusted analyses. Cannabis use was significantly associated with drug RAB scores in unadjusted analyses (mean difference 1.7 points) and analyses adjusted for age, sex, and employment (mean difference 1.3 points). When also adjusting for stimulant use, the impact of cannabis use was attenuated and no longer statistically significant (mean difference 1.1 points). Because of the central role of risky drug behaviors in the Russian HIV epidemic, it is important to understand how the use of multiple substances, including cannabis and alcohol, impacts risky drug behaviors.

Keywords Cannabis · Alcohol · Russia · HIV · Risk behaviors

A. Y. Walley (✉) · D. M. Cheng · C. Bridden · J. H. Samet
Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Department of Medicine, Boston Medical Center and Boston University School of Medicine, 91 E. Concord St., Suite 200, Boston, MA 02118, USA
e-mail: awalley@bu.edu

E. M. Krupitsky · V. Y. Egorova · E. E. Zvartau
St. Petersburg Scientific Research Center of Addictions and Psychopharmacology, St. Petersburg State Pavlov Medical University, St. Petersburg, Russia

D. M. Cheng
Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA

A. Raj · J. H. Samet
Department of Social and Behavioral Sciences, Boston University School of Public Health, Boston, MA, USA

E. M. Edwards
Data Coordinating Center, Boston University School of Public Health, Boston, MA, USA

G. E. Woody
Department of Psychiatry and Veterans Affairs Medical Center, University of Pennsylvania, Philadelphia, PA, USA

Introduction

Since the mid-1990s, HIV infection in the Russian Federation has become the largest HIV epidemic in Europe with an estimated 860,000 (range: 420,000–1.4 million) people living with HIV in 2003 (UNAIDS, 2005). Injection drug use has driven this rapid growth with more than 80% of officially reported HIV cases occurring among injection drug users (IDUs) (Dehne, Khodakevich, Hamers, & Schwartlander, 1999; Dehne, Pokrovskiy, Kobyshcha, & Schwartlander, 2000). Preventing HIV transmission in Russia is challenging because unsafe injection practices, such as sharing needles, are common (Shaboltas et al., 2006), and opioid replacement therapy with methadone or buprenorphine is illegal. Naltrexone is approved and was shown to prevent relapse to heroin use over a 6-month period in 42–44% of subjects, but it has not been widely used since its cost ($3.50/50 mg tablet) is prohibitive (Krupitsky et al., 2004b; Krupitsky et al., 2006). With HIV infection rates among IDUs as high as 30% (Kozlov et al., 2006; Krupitsky et al., 2004a; Shaboltas et al., 2006), the epidemic is crossing over into other risk groups, such as sex workers and sexual partners of injectors (UNAIDS, 2005).

Russian society has a long history of high levels of alcohol consumption. Nationally representative samples from 1992 to 2004 report about 70% of men and 45% of women drink alcohol (Zohoori, Banchette, & Popkin, 2005). Binge drinking is particularly popular in Russia, where the 20% of men with the greatest intake average more than 100 g (i.e., seven drinks) of alcohol per day (McKee, 1999). With an estimated one third of all deaths in Russia in 1994 related to alcohol, alcohol use is a significant Russian public health issue (Nemtsov, 2002; World Health Organization, 2004). Emerging evidence shows active alcohol use among HIV-infected patients in Russia and associations between heavy drinking and HIV risk behaviors (Benotsch et al., 2006; Krupitsky et al., 2005; Somlai et al., 2002).

Little has been published on the potential association between cannabis use and HIV risk behaviors, yet it is the most abused illicit drug in Russia (United Nations Office on Drugs and Crime, 2006). Among 15 and 16 year olds, 5% are monthly users, 22% are lifetime users and 24% report cannabis is easily available (Hibell et al., 2004). Use of other drugs, such as stimulants, has been linked to HIV transmission and risk behaviors in several populations (Buchanan et al., 2006; Lorvick, Martinez, Gee, & Kral, 2006; Molitor et al., 1999; Pechansky et al., 2006). A recent study among HIV-negative IDUs in St. Petersburg, Russia reports increased HIV seroconversion among IDUs who inject stimulants (Kozlov et al., 2006).

With a rapidly expanding HIV epidemic among IDUs, in the setting of high levels of cannabis and alcohol consumption, it is important to determine whether cannabis and alcohol use impacts risky drug behaviors. If HIV transmission is increased by cannabis or alcohol consumption, efforts to reduce cannabis or alcohol use may be effective HIV prevention strategies. We examined the association of cannabis use and alcohol use with HIV drug-risk behaviors in Russian drug users in this secondary analysis of baseline data from three randomized control trials conducted between 1999 and 2005.

Methods

Participants

We studied 404 patients undergoing treatment for heroin dependence in and around St. Petersburg. Between 1999 and 2002, 332 of these subjects were recruited into two randomized controlled trials of naltrexone for relapse prevention (Krupitsky et al., 2004b; Krupitsky et al., 2006). Trained physician research associates recruited these subjects during inpatient treatment at addiction hospitals affiliated with the St. Petersburg Pavlov State Medical University or by referral from community psychiatrists after outpatient detoxification. Eligibility criteria for these two trials were identical: a diagnosis of heroin dependence; age between 18 and 40 years; education at the high school level or above; at least one relative willing to support the subject's treatment and supervise adherence to study medications; a stable address with a working phone; no regular use of psychiatric medication; a negative pregnancy test and commitment to contraceptive use, if female; and abstinence from heroin and other substances for at least one week prior to study entry.

The remaining 72 subjects were the heroin dependent subgroup of the Russian PREVENT (Partnership to Reduce the Epidemic Via Engagement in Narcology Treatment) study. PREVENT was a randomized controlled trial of an HIV prevention intervention with the goal of reducing HIV sexual risk behaviors (Samet et al., 2006). PREVENT subjects were enrolled in 2004 and 2005. Trained physician research associates approached patients at the Leningrad Regional Center for Addictions and the Medical Narcology Rehabilitation Center after initial detoxification, and assessed eligibility, offered participation, and conducted assessments. Eligibility criteria included age 18 years and older; no alcohol or other abused substances for at least 48 h; reported unprotected anal or vaginal sex in the past 6 months; willingness to undergo HIV testing per standard narcology hospital protocols or previous diagnosis of HIV infection; and provision of reliable contact information (i.e., a home telephone number, an address within 150 km of St. Petersburg, and a friend or family contact).

In the current analysis, we combined these cohorts, as subjects were recruited from the same geographic area, under similar conditions, and with similar entry criteria and were assessed with similar baseline study instruments and methods. All subjects provided written informed consent prior to enrollment in the studies.

Measures

Baseline assessment in each cohort included the following: demographic survey; the Risk Assessment Battery (RAB) to assess HIV drug risk behavior in the past 6 months (Metzger et al., 1992; Metzger, Navaline, & Woody, 2001; Navaline et al., 1994); and the 30-day Timeline Follow-back (TLFB) survey for alcohol and cannabis use (Fals-Stewart, O'Farrell, Freitas, McFarlin, & Rutigliano, 2000; Sobell & Sobell, 1992; Sobell & Sobell, 1995). For the Naltrexone cohort, stimulant use was assessed using the 30-day TLFB. We categorized subjects with any stimulant use in the 30 days prior as stimulant users. For the PREVENT cohort, stimulant use in the last 30 days was assessed using the Addiction Severity Index (ASI) (McLellan et al., 1985). The ASI includes questions about

cocaine and amphetamine use in the past 30 days. Those subjects who answered yes to either of these questions were categorized as stimulant users. All study instruments were translated into Russian and checked for clarity by Russian research staff. Trained research staff delivered all survey instruments in one-on-one confidential sessions, except in the Russian PREVENT cohort, where the RAB was administered by an Audio Computer-Assisted Self-Interviewing (ACASI) System. ACASI allows additional privacy, minimizes literacy issues, encourages truth telling, and provides an identical recording of each question. Using this system has been shown to enhance the quality of self-report behavioral assessments and to provide an acceptable method for collecting self-reports of HIV risk behavior (Newman et al., 2002). HIV infection was recorded as part of the intake assessment for the PREVENT cohort, but not for the naltrexone cohorts. Subjects in all cohorts were compensated US$ 5 for participation in the baseline assessment.

Main Independent Variables

Heavy Alcohol Use

The TLFB survey asked subjects to recall the amount of alcohol consumed on each of the 30 days prior to entering treatment. From these data on their average weekly intake, alcohol status was classified into the following two categories: heavy (>seven drinks/week for women and >14/week for men) and not heavy (alcohol use less than heavy thresholds). We chose these categories to be consistent with National Institute on Alcohol Abuse and Alcoholism recommendations for heavy drinking. There are no established heavy drinking limits in Russia.

Cannabis Use

The TLFB survey in all cohorts asked subjects to recall cannabis use on each of the 30 days prior to entering treatment. We categorized subjects with any cannabis use in the 30 days prior as a cannabis user and those with no cannabis use as a cannabis non-user.

Outcomes

Measures of HIV Drug-Risk Behavior

The primary outcome of the study was the drug RAB score, modeled as a continuous variable. The drug RAB score ranges from 0 to 25. Injecting drugs within the last 6 months scored 1 point. Additional points were scored for needle sharing, equipment sharing, injecting in group settings, and mixing and sharing drugs. Among a prospectively followed group of IDUs in the United States, the drug RAB was able to discriminate those who seroconverted from those who remained seronegative for HIV infection (Metzger et al., 2001). A secondary outcome was risky drug use in the past 6 months. We defined risky drug use as having a drug RAB score greater than 1, which indicates some needle, equipment, or drug sharing in the last 6 months.

Data Analysis

We performed descriptive analyses (e.g. means, medians, standard deviations, proportions) for the following variables: age, sex, employment status, HIV status (available for PREVENT study only), stimulant use, cannabis use, heavy alcohol use, mean drug RAB score, and risky drug use. Subjects in the two naltrexone trials were grouped into the Naltrexone cohort ($n = 332$) and compared to the PREVENT cohort ($n = 72$) using two-sample t-tests or χ^2 tests. We assessed whether heavy alcohol use and cannabis use were associated with drug RAB score using multiple linear regression models, adjusting for age, sex, employment status, stimulant use and study cohort as covariates. Logistic regression models were used to analyze the secondary dichotomous outcome, risky drug use. To assess whether the effect of heavy alcohol use and cannabis use was similar for both cohorts, regression models were also fit separately for each cohort. For the PREVENT cohort, we performed an additional model that included a covariate for HIV status. All analyses were conducted using two-sided tests and a significance level of 0.05.

Results

Characteristics of the combined and individual cohorts are presented in Table 1. The combined cohort included 27% women, 28% employed, 13% stimulant users, 10% drinking at heavy levels, and 91% with a drug RAB score greater than 1. The Naltrexone and PREVENT cohorts were similar overall, but the PREVENT cohort was older (27 years vs. 23; $t(402) = 5.81$; $P < 0.01$); had more cannabis users (33% vs. 22%; χ^2 (1, $n = 404$) = 4.06; $P < 0.05$); and had a higher mean drug RAB score (12.6 vs. 8.0; $t(402) = 6.86$; $P < 0.01$). HIV infection status was only known for the PREVENT cohort, in which 35% were HIV-infected. The cohorts were similar with regards to gender, employment, stimulant use and heavy alcohol use. When stratified by heavy alcohol use, characteristics were similar for heavy and non-heavy alcohol users. When stratified by cannabis use, cannabis and non-cannabis users were similar on all characteristics except cannabis users were more likely to use stimulants (22% vs. 10%; χ^2 (1, $n = 404$) = 9.96; $P < 0.01$).

Table 1 Characteristics of IDUs in St. Petersburg overall and stratified by study cohort

	Total n = 404	PREVENT n = 72	Naltrexone n = 332	Test statistic[a]
Age, mean (SD)	23.9 (4.5)	26.6 (4.9)	23.4 (4.2)	5.81**
Male, no. (%)	294 (73)	49 (68)	245 (74)	0.98
Employed, no. (%)	111 (28)	18 (25)	93 (28)	0.27
HIV status, no. (%)				
Positive	25 (6)	25 (35)	0	NA
Negative	47 (12)	47 (65)	0	
Unknown	332 (82)	0	332 (100)	
Stimulant use, no. (%)	51 (13)	12 (17)	39 (12)	1.28
Cannabis use, no. (%)	97 (24)	24 (33)	73 (22)	4.06*
Heavy alcohol use no. (%)	40 (10)	7 (10)	33 (10)	0.0051
Risky IDU, no. (%)	368 (91)	67 (93)	301 (91)	0.42
Mean drug RAB score (SD)	8.9 (5.4)	12.6 (6.3)	8.0 (4.9)	6.86**

* $P < 0.05$, ** $P < 0.01$

[a] t-tests (degrees of freedom = 402) and χ^2 tests (degrees of freedom = 1)

Table 2 displays the bivariate analyses of cannabis use and heavy alcohol use with drug RAB score and risky drug use. In unadjusted analyses, cannabis users had a higher mean drug RAB score compared to non-users (10.2 vs. 8.5; $t(402) = 2.74$; $P < 0.01$), however there was no difference in the proportion of risky drug use (93% vs. 91%; χ^2 (1, n = 404) = 0.71; ns). We did not find a significant difference in drug RAB score between heavy and non-heavy drinkers (9.6 vs. 8.8; $t(402) = 0.89$; ns), however, heavy drinkers were more likely to be risky IDUs (100% versus 90%; Fisher's exact test; $P < 0.05$), as defined above in ''Outcomes''.

In multiple regression analyses, the mean drug RAB score remained higher for cannabis users compared to non-cannabis users (adjusted mean difference 1.09 points; $t(402) = 1.82$; ns; Table 3), however the magnitude of the association was attenuated and no longer statistically significant. Because cannabis use has been described as a gateway drug, and therefore may precede stimulant use in the causal pathway, we performed an adjusted analysis that did not include stimulant use which showed cannabis users had a mean drug RAB score 1.25 points higher than non-cannabis users ($t(402) = 2.10$; $P < 0.05$). No statistically significant association between heavy alcohol use and the continuous drug RAB score was detected in both the model adjusted for stimulant use (adjusted mean difference in drug RAB score 0.59 for heavy versus non-heavy alcohol use; $t(402) = 0.71$; ns) and the model that was not (adjusted mean difference 0.63; $t(402) = 0.76$; ns). Significant associations with lower drug RAB scores among covariates included older age, current employment, and participation in the Naltrexone cohort.

In secondary multiple linear regression analyses stratified by study cohort, associations between cannabis use and higher drug RAB scores were observed for both cohorts; however, the associations were not statistically significant. In both study cohorts, older age was significantly associated with lower drug RAB scores. Current employment was associated with lower RAB scores, though statistically significant for the PREVENT cohort, but not the Naltrexone cohort. In the PREVENT cohort analysis that included HIV status, subjects with HIV infection had higher drug RAB scores than those who were not HIV-infected (the mean drug RAB score was 4.6 points higher for HIV positive versus HIV negative subjects, $t(70) = 3.26$; $P < 0.01$).

Table 2 Bivariate comparisons of heavy alcohol use, cannabis use with drug risk behaviors

	Mean drug RAB score (SD)	T-test statistic (df = 400)	Risky IDU, no. (%)	χ^2 statistic (df = 1)
Heavy alcohol use	9.6 (5.3)	–0.89	40 (100)	*,[a]
No heavy alcohol use	8.8 (5.4)		326 (90)	
Cannabis use	10.2 (5.5)	2.74**	90 (93)	0.47
No cannabis use	8.5 (5.3)		276 (91)	

* $P < 0.05$, ** $P < 0.01$

[a] Calculated using Fisher's exact test

Table 3 Multiple linear regression models assessing the impact of heavy alcohol use and cannabis use on drug RAB score among IDUs in St. Petersburg (n = 404)

	Adjusted for stimulant use			Not adjusted for stimulant use		
	Mean change in drug RAB	Standard error	T-value	Mean change in drug RAB	Standard error	T-value
Heavy alcohol use	+0.59	0.84	0.71	+0.63	0.84	0.76
Cannabis use	+1.09	0.60	1.82	+1.25	0.59	2.10*
Age[a]	–0.23	0.059	–4.07**	–0.24	0.059	–4.15**
Female vs. Male	–0.30	0.57	–0.52	–0.23	0.57	–0.41
Employed	–1.47	0.56	–2.63**	–1.53	0.56	–2.74**
Naltrexone vs. PREVENT cohort	–5.13	0.68	–7.52**	–5.19	0.68	–7.60**
Stimulant use	+1.28	0.77	1.67	X	X	X

* $P < 0.05$, ** $P < 0.01$

[a] Adjusted mean corresponding to a 1 year increase in age

Unadjusted and adjusted logistic regression models showed no significant association between cannabis use and the secondary outcome risky drug use (unadjusted Odds Ratio 1.35; 95% confidence interval 0.57–3.19; χ^2 (1, n = 402) = 0.47; ns; adjusted Odds Ratio 1.30; 95% confidence interval 0.53–3.17; χ^2 (1, n = 402) = 0.32; ns). In logistic regression models including heavy alcohol use as an independent variable, odds ratios were not estimable as all heavy drinkers were risky drug users.

Discussion

Reducing drug-related HIV transmission in Russia will require wide adoption by the existing public health systems of effective prevention and treatment efforts, such as needle exchange programs, increased access to naltrexone and availability of methadone or buprenorphine replacement programs. Understanding the relationship between substance use (e.g., alcohol, cannabis, and stimulants) and injection drug use practices should inform the implementation of these programs. We hypothesized that among heroin dependent patients entering treatment in Russia, both heavy alcohol use and cannabis use would be independently associated with risky drug use behaviors. In unadjusted analysis, we found a statistically significant increase in drug RAB score among cannabis users. But in a multiple regression analysis this association was attenuated and not statistically significant after adjustment for stimulant use. We did not detect a significant association between heavy alcohol use and drug RAB score, yet we found that all heavy alcohol users had some needle, equipment, or drug sharing in the last 6 months.

How could cannabis use increase HIV drug risk behaviors? Research focused on the relationship of cannabis use and HIV drug risk behaviors is limited. Cannabis use is associated with the use of other drugs and frequently precedes the use of injectable drugs, such as heroin and cocaine (Fergusson, Boden, & Horwood, 2006; Golub & Johnson, 1994; Lynskey et al., 2003). A French cohort study of HIV-infected IDUs found that cessation of injection drug use was associated with decreased cannabis use (Bouhnik et al., 2004). No studies have directly assessed the hypothesis that cannabis use is associated with risky drug behaviors, either directly via immediate disinhibitory effects or as a marker of risky behavior.

Kozlov et al. reported increased incidence of HIV seroconversion among IDUs in St. Petersburg who use stimulants (Kozlov et al., 2006). Incidence was further increased with greater weekly injection frequency, however this study did not examine the impact of other drugs, such as alcohol or cannabis. It is possible that an association between heavy alcohol use, cannabis use and increased risk behaviors, as well as stimulant use and increased HIV seroconversion are examples of polysubstance use driving increased risk taking. In our adjusted model that did not include stimulant use, we found a significant association between cannabis use and increased drug RAB score, yet this association was attenuated and not significant in the model that included stimulant use. This attenuation may occur because cannabis use is a marker of one or more factors, such as stimulant use, which increases risk taking. Cannabis may be used concomitantly to attenuate the dysphoric results of stimulant use, or cannabis use may precede stimulant use in a causal pathway where cannabis use leads to stimulant use and then to increased risk taking.

The impact of cannabis use on HIV sex risk behaviors has been described (Brodbeck, Matter, &, Moggi, 2006; Simbayi et al., 2004; Somlai et al., 2002; Woody et al., 1999) but results have not always been inconsistent. Woody et al. (1999) found that cannabis use was not associated with increased sex risk among gay men, while

Brodbeck et al. (2006) found in a random sample of Swiss heterosexuals that cannabis use was associated with increased HIV sex risk behaviors, though not specifically at the time they were using cannabis. This is evidence that cannabis use is a marker of a riskier personality, or increases risk via chronic effects rather than directly increasing sex risk during the time it is used. In addition, chronic use may affect general risk taking by lowering motivation to protect oneself. If the effect of cannabis use is primarily an increase in general risk taking rather than acute situational risk taking, its use likely increases both HIV drug and sexual behaviors through a similar mechanism. HIV risk behavior research that addresses the impact of acute and chronic cannabis use on measures of motivation might further clarify the mechanism by which cannabis may increase risk behaviors.

Results of studies on alcohol's relationship to HIV risk among IDUs in the United States are mixed. Alcohol use among needle exchange participants in Providence, RI has been associated with increased injection and sexual risk (Stein et al., 2000). Using 30-day TLFB for both risky injection drug use and alcohol use, an association has been shown between daily alcohol use and daily risky injection (Stein, Charuvastra, Anderson, Sobota, & Friedmanna, 2002). Among Puerto Rican IDUs not in treatment, alcohol intoxication has been associated with sharing needles and cotton filters (Matos et al., 2004). However, among detoxification inpatients in Boston, alcohol consumption was associated with increased sexual risk, but not injection risk (Rees, Saitz, Horton, & Samet, 2001). Studies in Russia have shown that among HIV-infected inpatients, alcohol abuse or dependence has a significant association with sexual risk and a non-significant association with injection risk (Krupitsky et al., 2005). Among young injection drug users in St. Petersburg, Somlai et al. (2002) reported a 40% rate of needle sharing in the previous 90 days and a 74% rate of alcohol use in the previous 30 days, but this study did not look specifically at the association between alcohol and risky injection.

Alcohol and drug use disorders often co-occur in US populations (Belenko, 1979; Kessler et al., 1997), yet these Russian IDUs had a low percentage (10%) of heavy drinking. With the high overall prevalence of alcohol use in Russia, we expected to find higher rates of alcohol use among heroin users. The low proportion of heavy drinkers in these samples limited our ability to detect a significant association with risky drug behaviors. In unadjusted analyses, we found an association between heavy drinking and the secondary dichotomous outcome, risky drug use, but we were unable to estimate odds ratios from logistic regression models as all heavy drinkers also had risky drug behaviors. We did not find an association between heavy drinking and the continuous drug RAB score, however the measure of heavy drinking we used was developed in the United States and may not be the optimal measure in Russia. Because Russians consume more alcohol than Americans, it is conceivable that the optimal Russian measure may be a higher threshold, which in this study would have the impact of categorizing even fewer heavy drinkers. Despite the absence of a statistically significant association, we caution that alcohol may still represent an important predictor of HIV drug risk behaviors. As risky drug practices in Russia improve with more prevention education and optimal measures of heavy alcohol use in Russia are developed, a relationship between alcohol use and risky injection drug use may be revealed, as the heavy drinkers may be slow adopters of safer injection behaviors. Alternatively, the impact of heavy alcohol use may be limited to sexual risk behaviors.

The association of HIV infection with increased drug RAB scores observed in the PREVENT cohort is consistent with the rapid spread of HIV among IDUs in Russia. Furthermore, the mean drug RAB scores in both cohorts were higher than those seen among drug using Russian (Krupitsky et al., 2005) and American (Rees et al., 2001; Stein et al., 2000) cohorts and may well reflect increased risky drug behavior in Russia. The higher means among the PREVENT cohort relative to the Naltrexone cohort may be explained by the data collection methods. In PREVENT, drug risk behaviors were assessed with the use of the ACASI system which likely reduces social desirability bias and increases truth telling, whereas the Naltrexone cohort was assessed with one-on-one interviews, similar to the other studies that have used the RAB instrument.

There are several limitations to our study. First, our analysis was potentially underpowered to detect effects of the observed magnitude for cannabis and heavy alcohol use. Post-hoc calculations indicate that our study would have approximately 80% power to detect a minimum difference in drug RAB score of 2.5 and 1.8 for heavy drinking and cannabis use, respectively. Thus, it is likely that the study was not sufficiently powered to detect the observed magnitudes of association. Second, this is a cross-sectional analysis of baseline data, which limits our ability to establish causality, as well as the order of preceding causal elements. Third, this study combines results from three different cohorts that enrolled subjects at different times. However, the subjects were recruited and enrolled by similar study staff from similar settings and provided similar data. Fourth, while we did adjust for age, gender, and employment, future research should examine the impact of other psychosocial factors, such as mental illness on drug-related HIV risk behavior. Fifth, as all subjects came from the St. Petersburg area and were entering treatment, they may not adequately represent all Russian IDUs. IDUs entering treatment are likely to be more motivated to re-

duce not only their heroin use, but their alcohol and cannabis use as well. Therefore it is possible that they may have fewer drug-related risk behaviors before entering treatment. If this were the case, then this would decrease the likelihood we would find an association. Sixth, although we had cannabis use and heavy alcohol use data via the 30-day TLFB method, our risk behavior data was limited to the drug RAB score, which encompasses the prior 6 months. Collecting drug and alcohol use information as well as drug behavior information on a day-by-day basis would allow for examination of the daily relationship between substance use and risky behaviors.

Among 404 Russian IDUs entering treatment for heroin dependence, all heavy alcohol users had risky drug behaviors, but we did not find significant evidence of an association between heavy alcohol use and the HIV drug RAB score in unadjusted or adjusted analyses. We did find a significant association between cannabis use and drug RAB score after adjustment for age, sex, and employment status. Yet when adjusted for stimulant use, the association between cannabis use and HIV drug risk behaviors was attenuated and not statistically significant. Because of the central role of IDUs in the rapidly progressing Russian HIV epidemic, it is important to understand how the use of multiple substances, including cannabis and alcohol, impacts risky drug behaviors among IDUs.

Acknowledgments This work was supported in part by the National Institute on Alcohol Abuse and Alcoholism (NIAAA), NIH: R21-AA014821; NIDA U10-DA13043 and K05-DA 17009 (Dr. Woody). Dr. Walley was supported by National Institute on Drug Abuse (NIDA), NIH: R25-DA13582 and National Institute of Allergy and Infectious Diseases (NIAID), NIH: T32-AI52074 04. Dr. Samet is also supported by NIAAA: K24-AA015674. We would also like to acknowledge the assistance and support of Seville Meli and Suzette Levenson in the data collection and conceptualization of this project.

Reference

Belenko, S. (1979). Alcohol abuse by heroin addicts: Review of research findings and issues. *International Journal of Addictions, 14*, 965–975.

Benotsch, E. G., Pinkerton, S. D., Dyatlov, R. V., DiFranceisco, W., Smirnova, T. S., Dudko, V. Y., & Kozlov, A. (2006). HIV risk behavior in male and female Russian sexually transmitted disease clinic patients. *International Journal of Behavioral Medicine, 13*, 26–33.

Bouhnik, A. D., Carrieri, M. P., Rey, D., Spire, B., Gastaut, J. A., Gallais, H., & Obadia, Y. (2004). Drug injection cessation among HIV-infected injecting drug users. *Addictive Behaviors, 29*, 1189–1197.

Brodbeck, J., Matter, M., & Moggi, F. (2006). Association between cannabis use and sexual risk behavior among young heterosexual adults. *AIDS and Behavior, 10*, 599–605.

Buchanan, D., Tooze, J. A., Shaw, S., Kinzly, M., Heimer, R., & Singer, M. (2006). Demographic, HIV risk behavior, and health status characteristics of ''crack'' cocaine injectors compared to other injection drug users in three New England cities. *Drug and Alcohol Dependence, 81*, 221–229.

Dehne, K. L., Khodakevich, L., Hamers, F. F., & Schwartlander, B. (1999). The HIV/AIDS epidemic in eastern Europe: recent patterns and trends and their implications for policy-making. *AIDS, 13*, 741–749.

Dehne, K. L., Pokrovskiy, V., Kobyshcha, Y., & Schwartlander, B. (2000). Update on the epidemics of HIV and other sexually transmitted infections in the newly independent states of the former Soviet Union. *AIDS, 14*(Suppl 3), S75–S84.

Fals-Stewart, W., O'Farrell, T. J., Freitas, T. T., McFarlin, S. K., & Rutigliano, P. (2000). The timeline followback reports of psychoactive substance use by drug-abusing patients: psychometric properties. *Journal of Consulting and Clinical Psychology, 68*, 134–144.

Fergusson, D. M., Boden, J. M., & Horwood, L. J. (2006). Cannabis use and other illicit drug use: Testing the cannabis gateway hypothesis. *Addiction, 101*, 556–569.

Golub, A., & Johnson, B. D. (1994). The shifting importance of alcohol and marijuana as gateway substances among serious drug abusers. *Journal of Studies on Alcohol, 55*, 607–614.

Hibell, B., Andersson, B., Bjarnason, T., Ahlstrom, S., Balakireva, O., Kokkevi, A., & Morgan, M. (2004). The ESPAD Report 2003: Alcohol and other drug use among students in 35 European countries. Retrieved August 8, 2006, from http://www2.stakes.fi/english/publications_2000–2004/Kirjat%202004/m204.htm.

Kessler, R. C., Crum, R. M., Warner, L. A., Nelson, C. B., Schulenberg, J., & Anthony, J. C. (1997). Lifetime co-occurrence of DSM-III-R alcohol abuse and dependence with other psychiatric disorders in the National Comorbidity Survey. *Archives of General Psychiatry, 54*, 313–321.

Kozlov, A. P., Shaboltas, A. V., Toussova, O. V., Verevochkin, S. V., Masse, B. R., Perdue, T., Beauchamp, G., Sheldon, W., Miller, W. C., Heimer, R., Ryder, R. W., & Hoffman, I. F. (2006). HIV incidence and factors associated with HIV acquisition among injection drug users in St Petersburg, Russia. *AIDS, 20*, 901–906.

Krupitsky, E., Zvartau, E., Karandashova, G., Horton, N. J., Schoolwerth, K. R., Bryant, K., & Samet, J. H. (2004a). The onset of HIV infection in the Leningrad region of Russia: A focus on drug and alcohol dependence. *HIV Medicine, 5*, 30–33.

Krupitsky, E. M., Horton, N. J., Williams, E. C., Lioznov, D., Kuznetsova, M., Zvartau, E., & Samet, J. H. (2005). Alcohol use and HIV risk behaviors among HIV-infected hospitalized patients in St. Petersburg, Russia. *Drug and Alcohol Dependence, 79*, 251–256.

Krupitsky, E. M., Zvartau, E. E., Masalov, D. V., Tsoi, M. V., Burakov, A. M., Egorova, V. Y., Didenko, T. Y., Romanova, T. N., Ivanova, E. B., Bespalov, A. Y., Verbitskaya, E. V., Neznanov, N. G., Grinenko, A. Y., O'Brien, C. P., & Woody, G. E. (2004b). Naltrexone for heroin dependence treatment in St. Petersburg, Russia. *Journal of Substance Abuse Treatment, 26*, 285–294.

Krupitsky, E. M., Zvartau, E. E., Masalov, D. V., Tsoi, M. V., Burakov, A. M., Egorova, V. Y., Didenko, T. Y., Romanova, T. N., Ivanova, E. B., Bespalov, A. Y., Verbitskaya, E. V., Neznanov, N. G., Grinenko, A. Y., O'Brien, C. P., & Woody, G. E. (2006). Naltrexone with or without fluoxetine for preventing relapse heroin addiction treatment in St. Petersburg, Russia. *Journal of Substance Abuse Treatment, 31*, 319–328.

Lorvick, J., Martinez, A., Gee, L., & Kral, A. H. (2006). Sexual and injection risk among women who inject methamphetamine in San Francisco. *Journal of Urban Health, 83*, 497–505.

Lynskey, M. T., Heath, A. C., Bucholz, K. K., Slutske, W. S., Madden, P. A., Nelson, E. C., Statham, D. J., & Martin, N. G. (2003). Escalation of drug use in early-onset cannabis users vs

co-twin controls. *Journal of the American Medical Association, 289*, 427–433.

Matos, T. D., Robles, R. R., Sahai, H., Colon, H. M., Reyes, J. C., Marrero, C. A., Calderon, J. M., & Shepard, E. W. (2004). HIV risk behaviors and alcohol intoxication among injection drug users in Puerto Rico. *Drug and Alcohol Dependence, 76*, 229–234.

McKee, M. (1999). Alcohol in Russia. *Alcohol and Alcoholism, 34*, 824–829.

McLellan, A. T., Luborsky, L., Cacciola, J., Griffith, J., Evans, F., Barr, H. L., & O'Brien, C. P. (1985). New data from the Addiction Severity Index. Reliability and validity in three centers. *Journal of Nervous and Mental Disease, 173*, 412–423.

Metzger, D. S., DePhillips, D., Druley, P., O'Brien, C. P., McLellan, A. T., Williams, J., Navaline, H., Dyanick, S., & Woody, G. E. (1992). The impact of HIV testing on risk for AIDS behaviors. *NIDA Research Monograph, 119*, 297–298.

Metzger, D. S., Navaline, H. A., & Woody, G. E. (2001). Assessment of Substance Abuse: HIV Risk Assessment Battery. In D. Carson (Ed.), *Encyclopedia of drugs, alcohol and addictive behavior*. Farmington Mills, Michigan: Macmillan Reference USA.

Molitor, F., Ruiz, J. D., Flynn, N., Mikanda, J. N., Sun, R. K., & Anderson, R. (1999). Methamphetamine use and sexual and injection risk behaviors among out-of-treatment injection drug users. *American Journal of Drug and Alcohol Abuse, 25*, 475–493.

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.

Nemtsov, A. V. (2002). Alcohol-related human losses in Russia in the 1980s and 1990s. *Addiction, 97*, 1413–1425.

Newman, J. C., Des Jarlais, D. C., Turner, C. F., Gribble, J., Cooley, P., & Paone, D. (2002). The differential effects of face-to-face and computer interview modes. *American Journal of Public Health, 92*, 294–297.

Pechansky, F., Woody, G., Inciardi, J., Surratt, H., Kessler, F., Von Diemen, L., & Bumaguin, D. B. (2006). HIV seroprevalence among drug users: an analysis of selected variables based on 10 years of data collection in Porto Alegre, Brazil. *Drug and Alcohol Dependence, 82*(Suppl 1), S109–S113.

Rees, V., Saitz, R., Horton, N. J., & Samet, J. (2001). Association of alcohol consumption with HIV sex- and drug-risk behaviors among drug users. *Journal of Substance Abuse Treatment, 21*, 129–134.

Samet, J. H., Krupitsky, E., Cheng, D. M., Raj, A., Egorova, V., Levenson, S., Meli, S., Bridden, C., Verbitskaya, E., Kamb, M. L., & Zvartau, E. (2006, August). A randomized controlled trial (RCT) to reduce HIV sex-risk behaviors among hospitalized Russian substance dependent patients: The Russian Partnership to Reduce the Epidemic Via Engagement in Narcology Treatment (Russian PREVENT) Study. *Poster session presented at the XVI International AIDS Conference*, Toronto, Canada.

Shaboltas, A. V., Toussova, O. V., Hoffman, I. F., Heimer, R., Verevochkin, S. V., Ryder, R. W., Khoshnood, K., Perdue, T., Masse, B. R., & Kozlov, A. P. (2006). HIV prevalence, sociodemographic, and behavioral correlates and recruitment methods among injection drug users in St. Petersburg, Russia. *Journal of Acquired Immune Deficiency Syndromes, 41*, 657–663.

Simbayi, L. C., Kalichman, S. C., Jooste, S., Mathiti, V., Cain, D., & Cherry, C. (2004). Alcohol use and sexual risks for HIV infection among men and women receiving sexually transmitted infection clinic services in Cape Town, South Africa. *Journal of Studies on Alcohol, 65*, 434–442.

Sobell, L., & Sobell, M. (1992). Timeline followback: 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). Totawa, NJ: The Humana Press Inc.

Sobell, L., & Sobell, M. (1995). *Alcohol timeline followback (TLFB) Users' Manual*. Toronto, Canada: Addiction Research Foundation.

Somlai, A. M., Kelly, J. A., Benotsch, E., Gore-Felton, C., Ostrovski, D., McAuliffe, T., & Kozlov, A. P. (2002). Characteristics and predictors of HIV risk behaviors among injection-drug-using men and women in St. Petersburg, Russia. *AIDS Education and Prevention, 14*, 295–305.

Stein, M. D., Charuvastra, A., Anderson, B., Sobota, M., & Friedmanna, P. D. (2002). Alcohol and HIV risk taking among intravenous drug users. *Addictive Behaviors, 27*, 727–736.

Stein, M. D., Hanna, L., Natarajan, R., Clarke, J., Marisi, M., Sobota, M., & Rich, J. (2000). Alcohol use patterns predict high-risk HIV behaviors among active injection drug users. *Journal of Substance Abuse Treatment, 18*, 359–363.

UNAIDS (2005). *AIDS epidemic update: December 2005*. Retrieved August 8, 2006, from http://www.unaids.org/epi/2005/doc/EPI-update2005_pdf_en/epi-update2005_en.pdf.

United Nations Office on Drugs and Crime (2006). 2006 World drug report. Retrieved August 8, 2006, from www.unodc.org/unodc/world_drug_report.html.

World Health Organization (2004). Global status report on alcohol 2004: Russian Federation. Retrieved August 8, 2006, from http://www.who.int/substance_abuse/publications/alcohol/en/.

Woody, G. E., Donnell, D., Seage, G. R., Metzger, D., Marmor, M., Koblin, B. A., Buchbinder, S., Gross, M., Stone, B., & Judson, F. N. (1999). Non-injection substance use correlates with risky sex among men having sex with men: Data from HIVNET. *Drug and Alcohol Dependence, 53*, 197–205.

Zohoori, N., Banchette, D., & Popkin, B. (2005). Monitoring Health Conditions in the Russian Federation: The Russia Longitudinal Monitoring Survey 1992–2004. Report submitted to the U.S. Agency for international development. Retrieved August 8, 2006, from http://www.cpc.unc.edu/projects/rlms/papers.html.

The Patients in Recovery (PIR) Perspective: Teaching Physicians About Methamphetamine

Alexander Y. Walley, MD, MSc
Karran A. Phillips, MD, MSc
Adam J. Gordon, MD, MPH

ABSTRACT. Methamphetamine dependence is an emerging epidemic confronting physicians. In an effort to improve understanding of its impact, the authors presented an educational workshop at a national meeting for general internists featuring small group discussions with patients in recovery (PIR) from methamphetamine dependence. Participants rated the workshop highly, stating it would lead to concrete change in their teaching, research, or patient care practices and they would invite the workshop to their institution for presentation. Direct interaction with PIR was the most valued aspect of the workshop. Lessons learned included patient's fear of being "turned in" limits disclosure of methamphetamine use to physicians; active users have little insight into methamphetamine-related changes in physical appearance; and a sense of productivity reinforces ongoing methamphetamine use. Workshops that include small group discussions between physicians and PIR are an innovative, practical, and acceptable method to teach physicians about their role in helping patients with substance dependence.

KEYWORDS. Methamphetamine abuse, physician education, substance abuse training

INTRODUCTION

Substance use disorder (SUD) education has typically focused on the medical complications of late-stage alcohol dependence, probably because these complications are most commonly recognized in hospitalized patients where medical student and residency training has been focused (1). With the demonstrated benefit of screening and brief intervention techniques to

Alexander Y. Walley is affiliated with the Clinical Addictions Research and Education Unit, Section of General Internal Medicine, Boston Medical Center, Boston, Massachusetts, USA.

Karran A. Phillips is affiliated with the Division of General Internal Medicine, Johns Hopkins School of Medicine, Baltimore, Maryland, USA.

Adam J. Gordon is affiliated with the Mental Illness Research Education and Clinical Center, VISN 4, VA Pittsburgh Healthcare System, Pittsburgh, Pennsylvania, USA; Center for Health Equity Research and Promotion, VA Pittsburgh Healthcare System, Pittsburgh, Pennsylvania, USA; and Center for Research on Health Care, University of Pittsburgh, Pittsburgh, Pennsylvania, USA.

Address correspondence to: Alexander Y. Walley, MD, MSc, Boston Medical Center, 801 Massachusetts Avenue, 2nd Floor, Boston, MA 02218 USA (E-mail: awalley@bu.edu).

The authors would like to acknowledge the assistance of Deborah Service of the Matrix Institute and the 10 volunteer small group leaders at the workshop. The authors would also like to acknowledge Maryann Amodeo, MSW, PhD, for guidance in developing the workshop. Dr. Walley was supported by the National Institute on Drug Abuse (R25-DA13582) and National Institute of Allergy and Infectious Diseases (T32-AI52074). Dr. Gordon was supported by a VA HSR&D Research Career Development Award (RCD-00038–2). The results of this paper were presented at the 2006 Association for Medical Education and Research in Substance Abuse (AMERSA) national conference, Washington, DC, November 2007.

Substance Abuse, Vol. 29(4) 2008
Available online at http://www.haworthpress.com

doi: 10.1080/08897070802418493

reduce hazardous alcohol use in randomized controlled trials (2), SUD medical education has shifted from managing the late-stage complications of hospitalized patients towards preventing the progression and complications of alcohol problems (1). The emergence of the human immunodeficiency virus (HIV) epidemic and its association with injection drug use has increased attention on preventing and treating other SUDs, including methamphetamine dependence, an emerging problem that is challenging to treat and has significant medical, social, and environmental consequences (3).

Discussions between physicians and patients in recovery (PIR) from methamphetamine dependence may improve physicians' attitudes, knowledge, and skills in caring for all SUD patients. At a national meeting of general internists, we conducted a workshop that featured small group discussions with PIR from methamphetamine dependence focused on how physicians can effectively address methamphetamine dependence. We sought to determine if using PIR in small group discussions was practical, effective, and acceptable.

CASE

Our workshop, "Learning from the patient's perspective: Methamphetamine abuse," was presented at the Society for General Internal Medicine (SGIM) Annual Conference in April 2006 in Los Angeles, California. The learning objectives were to (1) describe the epidemiology and the acute and chronic health effects of methamphetamine; (2) summarize the available treatment options for methamphetamine-related disorders; (3) describe the experience and perspective of PIR; and (4) incorporate PIR as a method for teaching about addiction medicine. The 90-minute workshop began with two didactic presentations: the history, epidemiology, and physical effects of methamphetamine (20 minutes); and the treatment of methamphetamine-use disorders (15 minutes). We then divided into five small groups each with two PIR, and six to eight conference attendees to participate in a 40-minute discussion of methamphetamine use and treatment. After the small group discussions, a representative from each table shared lessons learned with the entire group (about 2 minutes per table). The final 5 minutes of the workshop were dedicated to the distribution, completion, and collection of evaluations.

Thirty-four SGIM conference attendees participated in our workshop. The participants included general academic internists and physicians-in-training. We had four facilitators: three internists with experience working with SUD patients and a local, nonphysician, SUD treatment specialist. The treatment specialist coordinated the participation of 10 volunteer PIR. Prior to the workshop, we asked the PIR to focus the discussions on the role physicians can play in the treatment of methamphetamine dependence. They were encouraged to let physicians' questions guide the discussions and avoid lengthy self-help group-style testimonials.

Lessons learned by the physician participants in the small group discussions included (1) patients' fear of being "turned in" to authorities can limit disclosure of methamphetamine use and its consequences to physicians; (2) active users have little insight into changes in physical appearance due to their use; and (3) in addition to euphoria, a sense of "productivity" among users reinforces ongoing use. In response to feedback solicited by the facilitators at the end of the workshop, the PIR expressed both surprise about the lack of knowledge among the physicians about methamphetamine and support for the PIR-physician interactive teaching format.

Of the 34 physician attendees, 24 (71%) returned evaluations compared to an overall conference evaluation return rate of 64% (Table 1). Our workshop received higher scores than the conference in general on all questions except Question 7 where workshop attendees indicated lower prior knowledge of methamphetamine and its treatment than other topics covered during the conference. Seventeen write-in comments regarding the workshop were generally positive, with praise for the inclusion of and interaction with PIR as the most common theme (Table 2).

TABLE 1. Workshop Evaluation

Question		Workshop average[a]	Conference average[a]
1	Overall session rating	4.83	4.31
2	Quality of content	4.83	4.31
3	Faculty presenters	4.57	4.38
4	Audience interaction	4.87	4.39
5	Inviting this workshop to your institution for presentation	4.58	3.72
6	Likelihood of concrete change in teaching, research, patient care or administration	4.21	3.65
7	Prior knowledge of topic	4.46	5.80
8	Audience size	2	2.01

[a]Questions 1 to 4 scale = 1 (poor) to 5 (outstanding).
Question 5 scale = 1 (no) to 5 (definitely).
Question 6 scale = 1 (definitely will not change) to 5 (definitely will change).
Question 7 scale = 1 (poor) to 10 (expert).
Question 8 was a 3-point scale with 1 = too small; 2 = optimal; 3 = too big.

TABLE 2. Selected Write-in Comments

- "Not only is this a fantastic way for doctors to learn from clients but this kind of contact approach has wider implications for reducing stigma. A huge compliment to you for taking the risk and involving clients."
- "Excellent to have people in recovery here—rare treat to have patients in a workshop. . .Would recommend putting interactive portion earlier or breaking up the lectures to engage participants earlier."
- "It was a privilege to talk to patients in recovery. This was definitely the highlight of the conference for me."
- "Nice combination of didactic with small group and the use of patients as educators."
- "Very much enjoyed having the clients there for expert opinion."

DISCUSSION

Directly involving PIR represents a novel means to educate physicians on the diagnosis, treatment, and management of SUDs. At a national conference of physicians, we found that utilizing PIR as interactive teachers was accepted and achieved high satisfaction as a teaching modality by the audience and PIR. PIR personalized the impact of methamphetamine, promoted the destigmatization of the disease, and clarified the role health care providers should play in caring for patients with SUDs.

Research demonstrates that only 51% of internal medicine residency programs had formal substance abuse curricula (4); large proportions of physicians neither screen, nor intervene or refer patients with SUDs to treatment (5,6); and professional satisfaction and confidence in treating patients with SUD are less than they are for treating patients with other common conditions, such as hypertension (7).

Curricula focused on screening and treatment for SUDs utilizing simulated patients can improve clinical skills for medical students (8), residents (9), and physicians (10). In addition, physician training that includes experiences with SUD treatment programs are associated with increased screening and referral to treatment (11). A patient-centered approach to care was recently promoted in the Institute of Medicine (IOM) report, "Improving the Quality of Health Care for Mental and Substance-Use Conditions"(12). A teaching modality where PIR are the teachers and physicians are the learners may be another patient-centered approach to augment learning of SUD.

Workshop participants reported a high likelihood that the lessons learned from PIR will result in concrete changes in their teaching, research, and patient care practices. The impact of this method on physician knowledge, skills, and behavior warrants further research. In conclusion, we found that focused small group discussions between PIR and physicians at a national conference workshop is an innovative, practical, and acceptable method to teach physicians about the role they can play in helping patients with addictions.

REFERENCES

1. Walsh RA: Medical education about alcohol: review of its role and effectiveness. Alcohol. 1995;30:689–702.

2. Whitlock EP, Polen MR, Green CA, Orleans T, Klein J: 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. Ann Intern Med 2004;140:557–568.

3. Rawson RA, Gonzales R, Brethen P: Treatment of methamphetamine use disorders: an update. J Subst Abuse Treat 2002;23:145–150.

4. . Isaacson JH, Fleming M, Kraus M, Kahn R, Mundt M: A national survey of training in substance use disorders in residency programs. J Stud Alcohol 2000;61:912–915.

5. 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.

6. Friedmann PD, McCullough D, Saitz R: Screening and intervention for illicit drug abuse: a national survey of primary care physicians and psychiatrists. Arch Intern Med 2001;161:248–251.

7. Saitz R, Friedmann PD, Sullivan LM, Winter MR, Lloyd-Travaglini C, Moskowitz MA, Samet JH: Professional satisfaction experienced when caring for substance-abusing patients: faculty and resident physician perspectives. J Gen Intern Med 2002;17:373–376.

8. Kahan M, Wilson L, Midmer D, Borsoi D, Martin D: Randomized controlled trial on the effects of a skills-based workshop on medical students' management of problem drinking and alcohol dependence Subst Abus 2003;24:5–16.

9. Parish SJ, Ramaswamy M, Stein MR, Kachur EK, Arnsten JH: Teaching about Substance Abuse with Objective Structured Clinical Exams. J Gen Intern Med 2006;21:453–459.

10. Saitz R, Sullivan LM, Samet JH: Training community-based clinicians in screening and brief intervention for substance abuse problems: translating evidence into practice. Subst Abus 2000;21:21–31.

11. Warburg MM, Cleary PD, Rohman M, Barnes HN, Aronson M, Delbanco TL: Residents' attitudes, knowledge, and behavior regarding diagnosis and treatment of alcoholism. J Med Educ 1987;62:497–>503.

12. Institute of Medicine. *Improving the Quality of Healthcare for Mental and Substance-Use Conditions.* Washington, DC: The National Academies Press; 2005.

Patient Navigation: State of the Art or Is it Science?

Kristen J. Wells, PhD, MPH[1]
Tracy A. Battaglia, MD, MPH[2]
Donald J. Dudley, MD[3]
Roland Garcia, PhD[4]
Amanda Greene, PhD, MPH, RN[5]
Elizabeth Calhoun, PhD[6]
Jeanne S. Mandelblatt, MD, MPH[7]
Electra D. Paskett, PhD[8]
Peter C. Raich, MD[9]
and the Patient Navigation Research Program

[1] Department of Health Outcomes and Behavior, Division of Cancer Prevention and Control, H. Lee Moffitt Cancer Center and Research Institute, Tampa, Florida.

[2] Women's Health Unit, Section of General Internal Medicine, Boston University School of Medicine, Boston, Massachusetts.

[3] Department of Obstetrics and Gynecology, Division of Maternal-Fetal Medicine, The University of Texas Health Science Center at San Antonio, San Antonio, Texas.

[4] Center to Reduce Cancer Health Disparities, National Cancer Institute, Bethesda, Maryland.

[5] NOVA Research Company, Bethesda, Maryland.

[6] Division of Health Policy and Administration, School of Public Health, University of Illinois at Chicago, Chicago, Illinois.

[7] Lombardi Comprehensive Cancer Center, Departments of Oncology and Medicine, Georgetown University Medical Center, Washington, District of Columbia.

[8] Division of Epidemiology, College of Public Health, The Ohio State University Comprehensive Cancer Center, Columbus, Ohio.

[9] Division of Hematology/Oncology, Department of Medicine, Denver Health and Hospital Authority, Denver, Colorado.

First implemented in 1990, patient navigation interventions are emerging today as an approach to reduce cancer disparities. However, there is lack of consensus about how patient navigation is defined, what patient navigators do, and what their qualifications should be. Little is known about the efficacy and cost-effectiveness of patient navigation. For this review, the authors conducted a qualitative synthesis of published literature on cancer patient navigation. By using the keywords 'navigator' or 'navigation' and 'cancer,' 45 articles were identified in the PubMed database and from reference searches that were published or in press through October 2007. Sixteen studies provided data on the efficacy of navigation in improving timeliness and receipt of cancer screening, diagnostic follow-up care, and treatment. Patient navigation services were defined and differentiated from other outreach services. Overall, there was evidence of some degree of efficacy for patient navigation in increasing participation in cancer screening and adherence to diagnostic follow-up care after the detection of an abnormality. The reported increases in screening ranged from 10.8% to 17.1%, and increases in adherence to diagnostic follow-up care ranged from 21% to 29.2% compared with control patients. There was less evidence regarding the efficacy of patient navigation in reducing either late-stage cancer diagnosis or delays in the initiation of cancer treatment or improving outcomes during cancer survivorship. There were methodological limitations in most studies, such as a lack of control groups, small sample sizes, and contamination with other interventions. Although cancer-related patient navigation interventions are being adopted increasingly across

This study was funded by the National Cancer Institute (NCI) through its Center to Reduce Cancer Health Disparities, National Institutes of Health, Department of Health and Human Services (UO1 CA 117281-01) and by the American Cancer Society (ACS) (SIRSG-05-253-01).

Dr. Wells' contribution to this article was funded by a grant from the NCI (R25 CA090314; Paul B. Jacobsen, PhD, principal investigator).

*Study findings are the result of the collaborative efforts of the following sites and principal investigators: Boston University Medical Center (Karen M. Freund, MD, MPH and Tracy A. Battaglia, MD, MPH); Chicago Patient Navigation Research Program (Charles Bennett, MD, PhD, MPP and Elizabeth Calhoun, PhD); Denver Health and Hospital Authority (Peter C. Raich, MD, FACP and Elizabeth Whitley, PhD, RN); George Washington University Cancer Institute, District of Columbia City-wide Patient Navigation Research Program (Steven R. Patierno, PhD); H. Lee Moffitt Cancer Center and Research Institute (Richard G. Roetzheim, MD, MSPH); Northwest Tribal Cancer Navigator Program (Victoria Warren-Mears, PhD); The Ohio State University Comprehensive Cancer Center (Electra D. Paskett, PhD); University of Rochester, Department of Family Medicine Randomized Controlled Trial of Primary Care-based Patient Navigation-Activation (Kevin Fiscella, MD, MPH); and University of Texas Health Science Center at San Antonio (Donald J. Dudley, MD) in conjunction with NOVA Research Company (Paul A. Young, MBA, MPH), the ACS, and the NCI Center to Reduce Cancer Health Disparities (Roland Garcia, PhD and Mary Ann Van Duyn, PhD, MPH, RD, Program Directors).

The following authors contributed substantially to the preparation of this article but were not included in the author listing because of the journal's limitation on the number of authors: Steven R. Patierno, PhD, GW Cancer Institute, George Washington School of Medicine and Health Sciences, George Washington School of Public Health and Health Services, Washington, DC; Richard G. Roetzheim, MD, MSPH, Department of Family Medicine, School of Medicine, University of South Florida, Tampa, Fla; and Paul A. Young, MBA, MPH, NOVA Research Company, Bethesda, Md.

Address for reprints: Kristen J. Wells, PhD, MPH, Moffitt Cancer Center, 12902 Magnolia Drive, FOW-EDU, Tampa, FL 33612; Fax: (813) 745-5525; E-mail: kristen.wells@moffitt.org

Received April 22, 2008; accepted May 27, 2008

DOI 10.1002/cncr.23815
Published online 8 September 2008 in Wiley InterScience (www.interscience.wiley.com).

the United States and Canada, further research will be necessary to evaluate their efficacy and cost-effectiveness in improving cancer care. ***Cancer* 2008;113:1999–2010.** © *2008 American Cancer Society.*

KEYWORDS: neoplasms, healthcare disparities, quality of healthcare, delivery of healthcare.

Continued advances in cancer screening and treatment are evident in the overall reduction in morbidity and mortality among those diagnosed with cancer.[1] Uninsured, African-American, Hispanic, and low-income patients are less likely than white, high-income, and insured patients to receive recommended cancer care.[2,3] Those most at risk for advanced stage at cancer diagnosis and high mortality include racial/ethnic minorities and socioeconomically disadvantaged populations who are more likely to be uninsured.[2] These same populations experience significant delays in completing follow-up care once a screening abnormality has been detected.[4,5] Although the reasons for these disparities are complex and are not understood completely, research has identified numerous patient, provider, and health system barriers for these at-risk populations.[6,7] A growing body of literature indicates that known barriers to care interfere with timely access to diagnosis and treatment once a screening abnormality has been identified.[8-18]

In an effort to reduce these disparities, patient navigation has been proposed as an innovative intervention to address known barriers to obtaining cancer care. Patient navigation is a model of care that is expanding rapidly in underserved communities and medical institutions across the nation.[19,20] However, despite the proliferation of patient navigator programs, there is little consensus about what constitutes patient navigation services, and there is little information on the efficacy of patient navigation in improving outcomes. The objectives of this review were to 1) describe the evolution of patient navigation as a model to address cancer disparities, 2) review current literature that defines patient navigation and its impact on cancer care, and 3) describe the goals of the Patient Navigation Research Program (PNRP), sponsored by the National Cancer Institute (NCI) and the American Cancer Society (ACS), as a means to address existing gaps in our knowledge regarding the efficacy of patient navigation. In synthesizing the literature, we sought to investigate the following questions: 1) What is patient navigation? 2) What do patient navigators do? 3) How is patient navigation distinct from existing cancer care services? 4) What are the qualifications of a patient navigator? 5) What are the target populations served by patient navigation programs? 6) What are the intended outcomes of patient navigation? 7) Where in the cancer care continuum do patient navigators provide services? and 8) What is the efficacy of patient navigation?

Historic Evolution of Patient Navigation in Cancer Care

The prevention, diagnosis, and treatment of cancer are complex processes that often require consultation with multiple medical specialists, in multiple settings, using numerous medical tests.[21] To understand the unique challenges faced by disadvantaged populations in accessing these complex processes, the ACS conducted a series of hearings in 1989 with low-income Americans throughout the United States. The results of those hearings were published in a report by the ACS entitled *Report to the Nation: Cancer in the Poor*.[6] It indicated that poor individuals face significant 'obstacles' to accessing cancer care services that prevent them from obtaining needed care, including: 1) widespread financial barriers, such as being unable to afford health insurance, Medicaid or Medicare ineligibility, losing employment that provides health insurance, and lack of affordable cancer services; 2) logistical barriers, such as a lack of transportation, living at a far geographic distance from healthcare, lack of reminder systems, and lack of understandable cancer information; and 3) sociocultural barriers, such as limited social support and inadequate health literacy.

In response to the results of the ACS report, Dr. Harold P. Freeman partnered with the ACS to create the first patient navigation program in Harlem, New York in 1990 targeting women with historically poor breast cancer outcomes.[21-23] This innovative program assisted low-income women in overcoming barriers to breast cancer screening and follow-up care. In addition to expanding screening and education services throughout the community, specified members of the community provided patient navigation services to women who had a clinical finding suspicious for cancer.[23]

Since the pioneering work of Dr. Freeman, there has been a growing commitment to support patient navigation services. In 2001, the President's Cancer Panel recommended that funding be provided to support community-based programs, such as patient

navigator programs, to assist individuals in obtaining 'cancer information, screening, treatment, and supportive services.'[24] Consequently, there has been an expansion in programs nationwide with funding from private foundations, including the ACS, the Avon Foundation, and the Susan B. Komen Breast Cancer Foundation,[19] as well as local, state, and federal governments and community organizations. In 2003, there were over 200 cancer care programs identified nationwide by the NCI that were providing patient navigation.[19] By 2007, the ACS funded more than 60 patient navigation programs across the United States.[25]

The federal government has made a substantial commitment to patient navigation through support of 3 separate programs. In 2005, NCI's Center to Reduce Cancer Health Disparities funded 8 sites for the PNRP (and, in collaboration with the ACS, a ninth site joined the PNRP). This program will test community-based navigation programs using a control group. In June 2005, the Patient Navigator, Outreach, and Chronic Disease Prevention Act of 2005 authorized federal grants to hire and train patient navigators to assist patients with cancer and other serious chronic diseases to obtain access to timely diagnostic, treatment, and follow-up care,[21] and $2.9 million was appropriated in 2007 for this initiative. In 2006, the Center for Medicare Services (CMS) funded 6 demonstration sites for pilot programs targeting minority Medicare beneficiaries with the objective of overcoming barriers in screening, diagnosis, and treatment of cancer.[26] Despite the interest in patient navigation and funding of these programs, there is limited published information regarding their efficacy and cost-effectiveness.

MATERIALS AND METHODS

Study Identification

The objective of the current literature review was to identify and summarize both descriptive and efficacy literature on patient navigation. A review of research literature in the National Library of Medicine was conducted in October 2007 by searching the PubMed database to identify articles describing patient navigation programs that were published at any time in English with human participants. Reference lists of identified articles also were reviewed for relevant publications. The inclusion criteria specified 1) published original articles and 2) a description of a patient navigator program related to cancer treatment, diagnosis, or screening. The PubMed database was searched using the keywords 'navigator' or 'navigation' and 'cancer.' The search produced 893 citations; however, when the abstracts of each article were reviewed, only 35 were related to cancer patient navigation.[23,27-60] An additional 7 studies were either found in reference lists from articles that were identified in the search or were included in the same journal issue as another published article.[19,21,22,61-64] Three additional articles were identified by study authors.[65-67] All 45 articles were reviewed, and any article that described a patient navigator program was retained for further analysis. Twenty-eight articles provided descriptive information on cancer patient navigator programs in sites across the United States and Canada. Of these articles, 16 provided information on outcomes of a patient navigation intervention (Table 1). These articles were used to provide descriptive information on patient navigation and evidence regarding its efficacy.

RESULTS

What Is Patient Navigation?

Several definitions of patient navigation have been published.[21,22,68,69] Although variations do exist, patient navigation generally is described as a barrier-focused intervention that has the following common characteristics:

- Patient navigation is provided to individual patients for a defined episode of cancer-related care (eg, evaluating an abnormal screening test).
- Although tracking patients over time is emphasized, patient navigation has a definite endpoint when the services provided are complete (eg, the patient achieves diagnostic resolution after a screening abnormality).
- Patient navigation targets a defined set of health services that are required to complete an episode of cancer-related care.
- Patient navigation services focus on the identification of individual patient-level barriers to accessing cancer care.
- Patient navigation aims to reduce delays in accessing the continuum of cancer care services, with an emphasis on timeliness of diagnosis and treatment and a reduction in the number of patients lost to follow-up.

What Do Patient Navigators Do?

Despite its narrow barrier-focused definition, patient navigation has been operationalized quite broadly in practice. The term 'navigator' has been applied to any type of service that assists individuals in overcoming obstacles from screening to treatment and in coping with challenges during survivorship. In our analysis of published articles

TABLE 1
Outcomes of Published Patient Navigation Efficacy Studies

Citation	Cancer	Design	Participants (Location)	Outcome Measures	Results
Cancer screening studies					
Dignan 2005[29]	Breast	Prospective RCT (face-to-face navigator intervention, telephone navigator intervention, control)	157 Native American women (Denver, Colo)	Adherence to mammography screening guidelines	Participants in either intervention group more likely to receive mammography according to guidelines after intervention than before intervention; telephone intervention more effective than face-to-face intervention
Fang 2007[35]	Cervical	Prospective comparison of cervical cancer screening intervention plus patient navigation or control group that received 2-h general health education session	Korean-American women; 50 in control group, 52 in intervention group	Difference between intervention and control in receipt of Pap screening at follow-up	39 of 52 intervention participants requested navigation services; intervention participants more likely to receive Pap smear than control participants ($P < .001$).
Jandorf 2005[45]	Colorectal	Prospective RCT (patient navigation or control)	40 participants in control group; 38 participants received patient navigation (East Harlem, NY)	Colorectal cancer screening adherence	1) At 3-mo chart review, more patient navigation participants scheduled endoscopy appointments ($P = .005$); and 2) at 6-mo chart review, more patient navigation patients had completed an endoscopy ($P < .02$)
Nash 2006[48]	Colorectal	Retrospective, comparison of patients who received care before and after patient navigator plus gastrointestinal suite improvement intervention	1767 patients who received diagnostic or screening colonoscopies either before or after intervention; patients who completed preadmission testing (Bronx, NY)	1) Rate of colonoscopies and 2) rate of broken appointments	1) Increase in the number of individuals who received screening colonoscopies and 2) broken appointment rate declined from 67.2% to 5.3%
Rahm 2007[52]	BRCA1/2 genetic counseling	Prospective RCT	125 participants referred for genetic counseling (Kaiser Permanente, Colo)	1) Genetic counseling participation within 9 mo of referral and 2) time from referral to completed genetic counseling appointment.	1) No significant difference in appointment attendance between navigation and usual care, not enough power to detect differences; and 2) patient navigator intervention participants had appointments scheduled significantly sooner than usual care participants
Tingen 1998[57]	Prostate	Prospective RCT; sites randomized to traditional prostate cancer education, peer-educator only, client-navigator only, or combination of peer-educator and client-navigator	1522 participants in a prostate cancer screening program (southeastern state)	Participation in free prostate cancer screening.	In multiple logistic regression, participants who received either client navigation intervention or combined intervention more likely to participate in screening program than prostate cancer education participants.
Weinrich 1998[58]	Prostate	Prospective RCT; sites randomized to traditional prostate cancer education, peer-educator only, client-navigator only, or combination of peer-educator and client-navigator.	1717 participants in a prostate cancer screening program (southeastern state)	Participation in free prostate cancer screening	African-American and total study participants who received either client navigation or peer education intervention were more likely to participate in screening program than traditional intervention participants; participants who received education alone were as likely to participate in screening as combined peer education and client navigation intervention participants

(*continued*)

TABLE 1
(*continued*)

Citation	Cancer	Design	Participants (Location)	Outcome Measures	Results
Cancer care after abnormality					
Battaglia 2007[27]	Breast	Retrospective comparison of women seen before and after navigation intervention	1332 women with abnormal screening (Boston, Mass)	Timely follow-up from referral to diagnostic resolution	Navigation participants more likely to have timely follow-up than participants screened before intervention. Intervention effect remained after 1) controlling for race, age, insurance status, reason for referral, and source of referral; and 2) using propensity score analysis to adjust for differences in pre- and postintervention samples
Ell 2002[32]	Breast	Prospective, study enrollees compared with nonenrollees; intervention included health education, navigation, and counseling	Women who received abnormal mammograms; 605 participants were compared with 695 nonenrollees (Los Angeles, Calif and New York, NY)	1) Adherence to follow-up care after abnormal mammogram, 2) timeliness of diagnostic resolution, and 3) timeliness of initiation of cancer treatment	1) Intervention participants more likely to adhere to follow-up recommendations than nonenrollees, 2) enrollees more likely to get to diagnostic resolution in a timely manner than nonenrollees, and 3) nonstatistically significant difference in timeliness of initiation of cancer treatment between enrollees and nonenrollees
Ell 2002[33]	Cervical	Prospective, study enrollees compared with nonenrollees; intervention included health education, navigation, and counseling	Women with low-grade and high-grade squamous intraepithelial lesions prescribed follow-up repeat screening; 196 women were enrolled in study and compared with 369 nonenrollees (Los Angeles, Calif)	Adherence to follow-up appointments	Intervention participants had significantly better rates of adherence to at least 1 follow-up appointment ($P = .0002$ and $P = .0001$).
Ell 2007[34]	Breast	Prospective RCT (patient navigation plus counseling or usual care)	Women who received abnormal mammograms; 96 in intervention group and 108 in control group (Los Angeles, Calif)	1) Adherence to diagnostic follow-up through diagnostic resolution, 2) timely adherence from index screen to diagnostic resolution, and 3) timely entry rates for cancer patients	1) Intervention group participants more likely to adhere to diagnostic follow-up than usual care participants or women who did not participate in study, 2) intervention group participants had more timely adherence than usual care participants and nonparticipants, and 3) intervention participants diagnosed with cancer were more likely to have timely entry rates (diagnosis, treatment) than usual care participants
Ferrante 2008[65]	Breast	Prospective RCT (usual care or usual care plus patient navigation)	Women with suspicious mammogram results (BI-RADS 4 or 5); 50 participants were assigned to usual care, and 55 participants were assigned to usual care plus patient navigation (Newark, NJ)	1) Time from abnormal mammogram to date of diagnostic resolution and 2) differences in anxiety and satisfaction between usual care and intervention groups	1) Mean diagnostic interval less in intervention group than usual care ($p = .001$); and 2) 1 mo after diagnostic resolution, anxiety lower and satisfaction higher in intervention group compared with usual care ($P < .001$)

(*continued*)

TABLE 1
(*continued*)

Citation	Cancer	Design	Participants (Location)	Outcome Measures	Results
Freeman 1995[37]	Breast, cervical, prostate, colorectal	Prospective; patients who received navigation compared with patients who did not receive navigation	1) Patients with an abnormal screening test for breast, cervical, prostate, or colorectal cancer (n=1136); and 2) patients with cancer (n=8) (Harlem, NY)	1) Whether participants obtained a biopsy following a suspicious/ abnormal finding and 2) amount of time to complete biopsy	1) Nonsignificant finding that 85.7% of navigated patients obtained a biopsy, whereas 56.5% of nonnavigated patients completed a biopsy; and 2) 71.4% of navigated patients completed biopsy in <4 wk, whereas 38.5% of nonnavigated patients completed the biopsy in <4 wk (P = .047)
Giese-Davis 2006[41]	Breast	Prospective, before and after comparison of navigation participants	29 women recently diagnosed with breast cancer (Santa Cruz, Calif)	Change over time (at baseline, 3 mo, 6 mo, and 9 mo) in depression, trauma symptoms, desire for information on breast cancer, emotional and social quality of life, self-efficacy to cope with cancer, and physician-patient relationship	Trauma symptoms and desire for breast cancer resource information decreased and emotional well being and cancer self-efficacy increased
Nash 2006[48]	Colorectal	Retrospective; comparison of patients who received care before and after patient navigator plus gastrointestinal suite improvement intervention	1767 patients who received diagnostic or screening colonoscopies either before or after intervention; patients who completed preadmission testing (Bronx, NY)	1) Rate of colonoscopies and 2) rate of broken appointments	1) Increase in the number of individuals who received screening colonoscopies and 2) broken appointment rate declined from 67.2% to 5.3%
Oluwole 2003[63]	Breast	Retrospective; comparison of patients who received care before and after intervention that included patient navigation, free cancer screening, and health education	12,480 patients who were seen after intervention implementation from January 1995 to December 2000; 324 patients diagnosed with breast cancer; comparison group received care from 1964 to 1986 (Harlem, NY)	1) Stage at diagnosis and 2) survival	1) Reduction in late-stage (III and IV) disease at presentation, from 49% before intervention to 21% after intervention (P <.001); 2) significant increase in early-stage diagnosis (0 and I), from 6% before intervention to 41% after intervention (P <.001); and 3) crude 5-y survival rate of patients treated after intervention was 70.2% compared with 39% 5-y survival rate of women with surgically treated cancer before intervention
Psooy, 2004[51]	Breast	Retrospective comparison of patients who received care before and after patient navigation program	536 patients who underwent core breast biopsy (Nova Scotia, Canada)	Time from screening abnormality to diagnostic resolution	Patient navigator intervention participants had significantly less time from screening abnormality to biopsy (P <.001)

RCT indicates randomized controlled trial; Pap, Papanicolaou stain; BI-RADS, Breast Imaging Reporting and Data System.

that describe patient navigation services, we identified 4 areas in which patient navigators frequently intervene: 1) overcoming health system barriers, 2) providing health education about cancer across the cancer continuum from prevention to treatment, 3) addressing patient barriers to cancer care, and 4) providing psychosocial support. To overcome health system barriers, patient navigators may coordinate cancer diagnostic or treatment care from multiple providers; assist patients with completing medical paperwork; schedule, confirm, reschedule, and attend appointments; and facilitate patient-provider communication.[27,28,31,34,36,37,40,45,48-53,55,56,58,59,61,70] When providing health education, patient navigators provide written information, discuss diagnostic and genetic tests, discuss treatment options, and answer patients' questions.[28,31,34,36,45,50-53,55,56,58,59,61] To overcome patient barriers to cancer care, a patient navigator may address issues such as lack of transportation, financial and insurance barriers, lack of childcare or language translation, low health literacy, or low literacy.[28,40,41,50,53,56,70] Patient navigators also provide psychosocial or emotional support, either directly or by referring patients to social workers or cancer support groups.[28,50,51,55,56]

How Is Patient Navigation Distinct From Existing Cancer Care Services?

Patient navigation shares characteristics with other models of patient assistance.[30] For instance, hospital-based social workers may provide health education materials to oncology patients; and community health workers, lay health advisors, or promotoras may promote cancer screening in the community. In addition, case management and patient advocate models often provide services similar to those provided by patient navigators, but these models also can differ distinctively.

Case managers work to assist the client in achieving optimal wellness, self-management, and functional capability by linking clients with service providers and resources throughout the continuum of health and human services and care settings.[71] Although the principles of case management (case identification, identifying barriers to care, developing individual plans to overcome barriers, tracking over time) are embedded in patient navigation, there are distinct differences. Most noteworthy, patient navigation focuses on 1 health condition instead of the broader objective of case management to improve health in general. In addition, patient navigation tends to track to the completion of a discrete set of health services instead of using long-term follow-up. Similar to patient navigators, a patient advocate helps resolve issues about healthcare, medical bills, and job discrimination related to a patient's medical condition.[72] However, the focus of patient advocates is on improving the healthcare system rather than delivering care to individual patients.[26] Although patient navigators also may perform tasks similar to those of patient advocates, their objective is to overcome individual and logistic barriers to the prevention, diagnosis, and treatment of a health concern.

What Are the Qualifications of a Patient Navigator?

Our review of the published literature yielded great variation in the personnel who provide patient navigation services. Patient navigation services frequently were provided by a lay patient navigator,[28,29,50,56] although several programs described navigators with undergraduate degrees,[32,33,65,67] master's degrees,[32,33] nurse practitioners or nurses,[55,58,59] social workers,[57,58] health educators,[59] clinic staff members,[49] research assistants,[45,52] and cancer survivors.[41,59] Typically, patient navigators are paid personnel rather than volunteers.

Although it appears that most patient navigators in the United States are receiving some training, the quality and training of that training are unclear. In an evaluation of patient navigation in Canada, virtually none of the patient navigators had received navigation training; however, the patient navigators who were evaluated were either nurses or social workers and, thus, had extensive knowledge of medicine or case management.[73]

What Are the Target Populations Served by Patient Navigation Programs?

Patient navigator programs that were reviewed serve mainly those populations most at risk for poor cancer outcomes. The programs in our review targeted many diverse and underserved populations in the United States, including inner-city residents,[23,27,32,33,37,40,45,56,63,65] Native Americans,[28,29,50,53] low-income populations,[43] minority populations,[35,49,57,58] and rural residents.[67] However, several patient navigation programs did not specifically target underserved populations.[31,36,41,51,52,55,59] For example, patient navigation has been provided to medical center patients[31,34,36,41,48,51,55,59,65] and patients in a managed care organization.[52]

What Are the Intended Outcomes of Patient Navigation?

Our review indicated that most patient navigation programs have been designed to improve the outcomes for cancer in a single, specific site of the body, such as breast cancer. By far, most programs described in the published literature focus on improving outcomes for

breast cancer.[27-29,31,32,34,40,41,51,63,65] Other patient navigation programs target cervical cancer,[33,35,49] colorectal cancer,[45,48] prostate cancer,[57,58] lung cancer,[55] and head and neck cancer.[36] Only 5 programs reported navigation programs that targeted multiple cancer sites.[37,43,50,59,67]

Where in the Cancer Care Continuum Do Patient Navigators Provide Services?

Patient navigation services also target improving cancer outcomes across the cancer care continuum. Several programs were implemented to increase screening,[28,29,35,45,48,49,57,58] improve follow-up care after an abnormal cancer screen,[27,28,32-34,37,40,43,48,51,65] reduce time from diagnosis to treatment of cancer,[55] improve cancer treatment and the psychosocial experience of cancer treatment,[34,36,40,41,56,59,67] and improve accrual and retention in clinical trials.[50,56] Less frequently, patient navigators have provided healthcare,[55] assisted in accrual and retention of clinical trial participants,[28] recruited individuals for cancer screening,[28] and sought to increase compliance with referrals to BRCA1/2 genetic testing.[52] To date, no published study has evaluated the efficacy of a patient navigation intervention for cancer survivors.

What Is the Efficacy of Patient Navigation?

Sixteen studies evaluated the efficacy of a patient navigation intervention using several different study designs (Table 1), all with different outcomes. Most studies focused on the receipt of cancer diagnostic care and treatment services. Although the majority of these 16 published studies targeted improving outcomes in diagnostic breast health services,[27,29,32,34,41,51,63,65] none of the studies reviewed had comparable outcomes. Most published studies used prospective designs comparing participants who had received patient navigation with patients who did not receive navigation.[29,33-35,37,41,45,52,57,58,65] Seven studies (43.8%) randomly assigned participants or clinics to a patient navigation intervention or a comparison group.[29,34,45,52,57,58,65] Two studies were limited by low sample sizes.[41,52]

Improving screening rates

Six published articles provided evidence of the efficacy of patient navigation in improving screening rates for 3 cancers.[29,35,45,48,57,58] The improvement in the rate of adherence to screening ranged from 10.8% to 17.1% when patients in a navigation group were compared with a control group. Limitations in the research designs precluded reaching definite conclusions regarding efficacy. Some articles reported that patient navigation was combined with educational outreach,[57,58] included in a multifaceted cognitive-behavioral intervention,[35] or combined with improvements in the hospital's gastrointestinal suite,[48] making it difficult to determine whether patient navigation alone significantly increased screening rates. In 2 studies,[29,45] participants were randomized to the patient navigation intervention or a control arm, whereas another 2 studies[57,58] compared patient navigation with other educational interventions.

Improving adherence to diagnostic services after an abnormality is detected

Several published articles reported that patient navigation resulted in improvements both in adherence to follow-up visits after the detection of a screening abnormality (improvements ranged from 21% to 29.2% when patient navigation was compared with a control group) and in the timeliness of obtaining care from screening abnormality to diagnostic resolution among patients who were screened for breast, cervical, prostate, and colorectal cancer.[27,32-34,37,48,51,65] Only 2 studies randomly assigned patients to a patient navigation intervention or a usual care group.[34,65] Other studies used historic comparisons or study nonparticipants.[27,32,33,37,48,51,63] In addition, 3 studies combined patient navigation with counseling,[32-34] making it difficult to determine whether improvements in follow-up care were because of patient navigation or because of more intense psychosocial intervention.

Stage of cancer diagnosis

The only study that examined the effect of patient navigation on disease stage at the time of cancer diagnosis reported reductions in late-stage cancer diagnosis associated with an intervention that included patient navigation, free cancer screening, and culturally sensitive health education.[63] Because that study involved a multimodal intervention, it is impossible to draw conclusions regarding the effect of the patient navigator intervention alone on disease stage at the time of diagnosis.

Improving cancer treatment

Information regarding the impact of patient navigation on the timeliness of initiating cancer treatment is mixed. One study reported no significant improvement in the timeliness of initiating breast cancer treatment for patients who received patient navigation and counseling compared with nonparticipants,[32] whereas another study reported that patients who received patient navigation and counseling had faster initiation of breast cancer treatment than par-

ticipants who were randomized to receive usual care.[34] The information obtained from both of these studies was limited, because the patient navigation intervention was combined with other services.

Implications of existing research

Despite the flurry of interest and large financial investment in implementing patient navigation programs nationally, there remains only limited evidence of their efficacy as a means to reduce cancer health disparities. To convert these demonstration projects into long-lasting public policy, scientifically rigorous efficacy data are needed to demonstrate the benefits of patient navigation. To achieve this objective, there needs to be standardization in the definition of patient navigation, including its tasks, target population, and intended outcomes. Standard metrics to assess the benefit of existing programs are paramount, and high-quality training should be provided to patient navigators. To date, the PNRP is the only large-scale research study to examine the effectiveness of patient navigation.

The Patient Navigation Research Program

The PNRP, a 5-year multisite clinical trial, is designed to provide data regarding the efficacy and cost-effectiveness of the patient navigation intervention model.[74] Eight academic research institutions and 1 health board that serves underserved populations were awarded funding in 2005. The PNRP defines patient navigation as support and guidance offered to individuals who have an abnormal cancer screening test or a cancer diagnosis with the objective of accessing the cancer care system and overcoming barriers to timely, quality care. In the PNRP, patient navigation targets those who are most at risk for delays in care, including racial and ethnic minorities, patients from low-income populations, uninsured patients, and patients from rural areas who have an abnormal cancer screening test for breast, cervical, colorectal, or prostate cancer. PNRP patient navigators identify individual barriers to care and then work with the healthcare team and other community agencies to assist patients in overcoming those barriers.

Patient navigator training

The PNRP National Patient Navigator Training and Education Committee provides fundamental or core training across all sites.[75] The national training is supplemented by local training at each of the sites. This committee has successfully implemented 2 in-person training sessions for over 250 patient navigators from the PNRP, ACS, and CMS patient navigation programs. The training sessions were implemented using multiple adult learning modalities, including traditional lecture, interactive formats, and role play with case scenarios. The training curriculum covered topics such as an overview of cancer, cancer screening, cancer treatment, communication, culture and diversity, barriers to care, and mapping resources. The efficacy of training was evaluated by using tests before and after training that were developed by the training committee. Continuing education occurs through regular Webinar training sessions and annual in-person sessions.

Evaluation of the PNRP study

The PNRP used a committee structure to define common data elements and clinical definitions to measure common outcomes consistently across sites. Ultimate outcomes of navigation are to reduce morbidity and mortality of cancer. PNRP focuses on the measurement of the following intermediate outcomes:

- time from a cancer-related abnormal screening finding to a definitive diagnosis of cancer or resolution of abnormality for those who do not have cancer;
- time from cancer diagnosis to initiation of cancer treatment;
- time from initiation to completion of primary cancer care for patients newly diagnosed with cancer;
- patient satisfaction with cancer care; and
- cost-effectiveness of patient navigation.

In evaluating the effectiveness of patient navigation, the objectives of the PNRP—timely diagnosis and treatment and ease of interacting with the medical care system—appear to have intrinsic value. The assumption underlying investments in navigation is that these costs will be offset by reductions in mortality as a result of the more timely resolution of an abnormal cancer screening test than would occur in the absence of navigation. This assumption will be correct if navigation moves an individual to an earlier stage at diagnosis (or a significantly smaller tumor) than would occur in the absence of navigation. Among individuals who may delay but will attend follow-up, stage shift may not be as dramatic or may not occur at all, depending on the length of delay relative to tumor growth. For instance, recent studies report delays of 25 days versus 42 days in diagnostic follow-up with and without navigation, respectively.[65] This 17-day delay, although statistically significant, will not affect stage at diagnosis.

Because navigation in the PNRP is focused on populations that historically are under screened,

navigators will be helping individuals with prevalent tumors that have more advanced stage than are seen in a regularly screened population. In this situation, mortality benefits from stage shift can be expected to be minimal. Navigation could improve mortality for these prevalent cases (and others with minimal delay) only if they improve adherence to effective treatments in populations that otherwise would not comply with therapy. Treatment navigation programs are being developed,[76] and treatment adherence aspects are being addressed by several of the PNRP sites, but their efficacy for improving survival has not been tested.

DISCUSSION

Despite gains in cancer screening, diagnosis, and treatment, certain populations continue to suffer poor outcomes and higher mortality.[3,77] Several health system and individual barriers exist for underserved populations in accessing and completing recommended cancer care.[6,7] Although they originally were designed to overcome the barriers experienced by underserved patients who had screening abnormalities, today, patient navigation service programs are widespread throughout the United States and Canada and target several different cancer-related outcomes in many populations.

Sixteen published articles provided data on the efficacy of patient navigation. Some published articles indicate that patient navigation is associated with improvements in breast, prostate, and colorectal cancer screening as well as improvements in adherence to follow-up visits after the detection of an abnormality and reduction in the time from an abnormal screening to diagnostic resolution for breast, cervical, prostate, and colorectal cancer.[27,29,32-34,37,45,48,51,57,58,65] However, published studies have limitations that preclude drawing definitive conclusions about the efficacy of patient navigation, such as lack of a control group, lack of randomization to treatment or comparison groups, low sample size, no single definition of patient navigation, and combining patient navigation services with other interventions. Thus, information about the efficacy of patient navigation programs is limited, and there is no information about cost-effectiveness of patient navigation.

The PNRP is a collaborative effort that was designed to overcome limitations in the research literature by evaluating the efficacy and cost-effectiveness of various models of patient navigation in cancer in a standardized, rigorous process. The PNRP is collecting standardized data across 9 sites to determine the characteristics of successful and cost-effective navigation programs and to identify which patients benefit from patient navigation. Until more information is available regarding the efficacy of patient navigation programs, institutions that are considering implementing such interventions should be aware of the paucity of data regarding such benefits.

To date, there are no formal, recognized certification programs for patient navigators, nor is there evidence indicating which characteristics of navigators are most efficacious in improving cancer outcomes. Patient navigators have a variety of backgrounds and levels of formal education, and there is little information regarding the training of navigators.[21,73] The PNRP program has provided centralized, standardized training to patient navigators from all 9 sites and is collecting data to determine the characteristics of navigators who predict better outcomes in abnormal screening tests as well as cancer diagnosis and treatment. If it is determined that patient navigation is effective, then it will be important to agree on a standardized navigator training program and evaluate the appropriateness of a formal certification process for patient navigators. Without such formal certification, it will be difficult to obtain reimbursement from insurance companies or the federal government for patient navigation services.

In conclusion, patient navigation is an intervention designed to reduce health disparities by addressing specific barriers to obtaining timely, quality healthcare. This intervention is used in many different settings to target various cancer outcomes in many different populations. Although published research indicates that patient navigation may be associated with improvements in screening and diagnostic resolution after screening in certain populations, the research limitations preclude drawing generalizable conclusions regarding efficacy of patient navigation. A thorough evaluation of the PNRP and other scientifically rigorous, future programs will be necessary to ensure that navigator programs are effective and cost-effective before continued dissemination.

REFERENCES

1. Jemal A, Siegel R, Ward E, Murray F, Xu J, Thun MJ. Cancer statistics 2007. *CA Cancer J Clin*. 2007;57:43-66.
2. National Cancer Policy Board. Hewitt M, Simone JV, eds. Ensuring Quality Cancer Care. Institute of Medicine and Commission on Life Sciences, National Research Council. Washington, DC: National Academy Press; 1999.
3. The Commonwealth Fund. The Commonwealth Fund Commission on a High Performance Health System, Why Not the Best? Results from a National Scorecard on US Health System Performance. The Commonwealth Fund, September 2006. Available at: http://www.commonwealthfund.org/publications/publications_show.htm?doc_id=401577. Accessed October 8, 2007.

4. Shavers VL, Brown ML. Racial and ethnic disparities in the receipt of cancer treatment. *J Natl Cancer Inst*. 2002;94: 334-357.
5. Elmore JG, Nakano CY, Linden HM, Reisch LM, Ayanian JZ, Larson EB. Racial inequalities in the timing of breast cancer detection, diagnosis and initiation of treatment. *Med Care*. 2005;43:141-148.
6. American Cancer Society. Cancer and the Poor: A Report to the Nation. Atlanta, Ga: American Cancer Society; 1989.
7. Mandelblatt JS, Yabroff KR, Kerner JF. Equitable access to cancer care services: a review of barriers to quality cancer care. *Cancer*. 1999;86:2378-2390.
8. Lannin DR, Mathews HF, Mitchell J, et al. Influence of socioeconomic and cultural factors on racial differences in late-stage presentation of breast cancer. *JAMA*. 1998;279: 1801-1807.
9. 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.
10. Haas JS, Phillips KA, Sonneborne D, et al. Effect of managed care insurance on the use of preventive care for specific ethnic groups in the United States. *Med Care*. 2002;40:743-751.
11. Lee-Feldstein A, Feldstein PJ, Buchmueller T, Katterhagen G. The relationship of HMOs, health insurance, and delivery systems to breast cancer outcomes. *Med Care*. 2000;38: 705-718.
12. Rojas M, Mandelblatt J, Cagney K, Kerner J, Freeman H. Barriers to follow-up of abnormal screening mammograms among low-income minority women. Cancer Control Center of Harlem. *Ethn Health*. 1996;1:221-228.
13. Perez-Stable EJ, Sabogal F, Otero-Sabogal R, Hiatt RA, McPhee SJ. Misconceptions about cancer among Latinos and Anglos. *JAMA*. 1992;268:3219-3223.
14. Royak-Schaler R, deVellis BM, Sorenson JR, Wilson KR, Lannin DR, Emerson JA. Breast cancer in African-American families. Risk perception, cancer worry, and screening practices of first-degree relatives. *Ann N Y Acad Sci*. 1995;768:281-285.
15. Friedman LC, Webb JA, Weinberg AD, Lane M, Cooper HP, Woodruff A. Breast cancer screening: racial/ethnic differences in behaviors and beliefs. *J Cancer Educ*. 1995;10:213-216.
16. Hughes C, Lerman C, Lustbader E. Ethnic differences in risk perception among women at increased risk for breast cancer. *Breast Cancer Res Treat*. 1996;40:25-35.
17. Woloshin S, Schwartz LM, Katz SJ, Welch HG. Is language a barrier to the use of preventive services? *J Gen Intern Med*. 1997;12:472-477.
18. McCarthy BD, Yood MU, Janz NK, Boohaker EA, Ward RE, Johnson CC. Evaluation of factors potentially associated with inadequate follow-up of mammographic abnormalities. *Cancer*. 1996;77:2070-2076.
19. Hede K. Agencies look to patient navigators to reduce cancer care disparities. *J Natl Cancer Inst*. 2006;98:157-159.
20. Capitman JA, Bhalotra S, Ruwe M. Cancer and Elders of Color: Opportunities For Reducing Health Disparities. London, United Kingdom: Ashgate Publishing; 2005.
21. Newman-Horm PA,C-Change. Cancer Patient Navigation: Published Information. Washington, DC: C-Change; 2005.
22. Freeman HP. A model patient navigator program. *Oncol Issues*. 2004;19:44-46.
23. Freeman HP, Chu KC. Determinants of cancer disparities: barriers to cancer screening, diagnosis, and treatment. *Surg Oncol Clin North Am*. 2005;14:655-669.
24. Freeman HP. Voices of a Broken System: Real People, Real Problems. President's Cancer Panel Report of the Chairman 2000-2001. Bethesda, Md: National Institutes of Health; 2001.
25. American Cancer Society. Cancer Patients Can More Easily Navigate Health Care System Thanks to the American Cancer Society and AstraZeneca. Atlanta, Ga: American Cancer Society; 2008. Available at: http://www.cancer.org/docroot/MED/content/MED_2_1x_Cancer_Patients_Can_More_Easily_Navigate_Health_Care_System_Thanks_to_the_American_Cancer_Society_and_AstraZeneca.asp?sitearea=MED. Accessed April 10, 2008.
26. Institute for Alternative Futures. Patient Navigator Program Overview April 13, 2007. A Report for the Disparity Reducing Advances (DRA) Project. Available at: http://www.altfutures.com/DRA/Patient%20Navigator%20Overview.pdf. Accessed December 28, 2007.
27. 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):359-367.
28. Burhansstipanov L, Wound DB, Capelouto N, et al. Culturally relevant "navigator" patient support. The Native sisters. *Cancer Pract*. 1998;6:191-194.
29. Dignan MB, Burhansstipanov L, Hariton J, et al. A comparison of 2 Native American navigator formats: face-to-face and telephone. *Cancer Control*. 2005;12(suppl 2):28-33.
30. Dohan D, Schrag D. Using navigators to improve care of underserved patients: current practices and approaches. *Cancer*. 2005;104:848-855.
31. Doll R, Stephen J, Barroetavena MC, et al. Patient navigation in cancer care: program delivery and research in British Columbia. *Can Oncol Nurs J*. 2003;13:193-193.
32. Ell K, Padgett D, Vourlekis B, et al. Abnormal mammogram follow-up: a pilot study women with low income. *Cancer Pract*. 2002;10:130-138.
33. Ell K, Vourlekis B, Muderspach L, et al. Abnormal cervical screen follow-up among low-income Latinas: project SAFe. *J Womens Health Gend Based Med*. 2002;11:639-651.
34. Ell K, Vourlekis B, Lee P-J, Xie B. Patient navigation and case management following an abnormal mammogram: a randomized clinical trial. *Prev Med*. 2007;44:26-33.
35. Fang CY, Ma GX, Tan Y, Chi N. A multifaceted intervention to increase cervical cancer screening among underserved Korean women. *Cancer Epidemiol Biomarkers Prev*. 2007;16: 1298-1302.
36. Fillion L, de Serres M, Lapointe-Goupil R, et al. Implementing the role of a patient-navigator nurse at a university hospital centre. *Can Oncol Nurs J*. 2006;16:5–17.
37. Freeman HP, Muth BJ, Kerner JF. Expanding access to cancer screening and clinical follow-up among the medically underserved. *Cancer Pract*. 1995;3:19-30.
38. Freeman HP. Patient navigation: a community centered approach to reducing cancer mortality. *J Cancer Educ*. 2006; 21(suppl.):S11-S14.
39. Freeman H. Patient navigation: a community based strategy to reduce cancer disparities. *J Urban Health*. 2006;83: 139-141.
40. Frelix GD, Rosenblatt R, Solomon M, Vikram B. Breast cancer screening in underserved women in the Bronx. *J Natl Med Assoc*. 1999;91:195-200.
41. Giese-Davis J, Bliss-Isberg C, Carson K, et al. The effect of peer counseling on quality of life following diagnosis of breast cancer: an observational study. *Psycho-oncology*. 2006; 15:1014-1022.
42. Henry-Tillman R, Deloney LA, Savidge M, Graham CJ, Klimberg VS. The medical student as patient navigator as an approach to teaching empathy. *Am J Surg*. 2002;183: 659-662.

43. Hiatt RA, Pasick RJ, Stewart S, et al. Community-based cancer screening for underserved women: design and baseline findings from the Breast and Cervical Cancer Intervention Study. *Prev Med*. 2001;33:190-203.
44. Jandorf L, Fatone A, Borker PV, et al. Creating alliances to improve cancer prevention and detection among urban medically underserved minority groups: the East Harlem Partnership for Cancer Awareness. *Cancer*. 2006;107(8 suppl):2043-2051.
45. Jandorf L, Gutierrez Y, Lopez J, Christie J, Itzkowitz SH. Use of a patient navigator to increase colorectal cancer screening in an urban neighborhood health clinic. *J Urban Health*. 2005;82:216-224.
46. Lantz PM, Keeton K, Romano L, Degroff A. Case management in public health screening programs: the experience of the National Breast and Cervical Cancer Early Detection Program. *J Public Health Manag Pract*. 2004;10:545-555.
47. Lemak CH, Johnson C, Goodrick EE. Collaboration to improve services for the uninsured: exploring the concept of health navigators as interorganizational integrators. *Health Care Manag Rev*. 2004;29:196-206.
48. Nash D, Azeez S, Vlahov D, Schori M. Evaluation of an intervention to increase screening colonoscopy in an urban public hospital setting. *J Urban Health*. 2006;8:231-243.
49. Nguyen TT, McPhee SJ, Bui-Tong N, et al. Community-based participatory research increases cervical cancer screening among Vietnamese-Americans. *J Health Care Poor Underserved*. 2006;17:31-54.
50. Petereit DG, Rogers D, Burhansstipanov L, et al. Walking forward: the South Dakota Native American Project. *J Cancer Educ*. 2005;20(1 suppl):65-70.
51. Psooy BJ, Schreuer D, Borgaonkar J, Caines JS. Patient navigation: improving timeliness in the diagnosis of breast abnormalities. *Can Assoc Radiol J*. 2004;55:145-150.
52. Rahm AK, Sukhanova A, Ellis J, Mouchawar J. Increasing utilization of cancer genetic counseling services using a patient navigation model. *J Genet Counsel*. 2007;16:171-177.
53. Rogers D, Petereit DG. Cancer disparities research partnership in Lakota Country: clinical trials, patient services, and community education for the Oglala, Rosebud, and Cheyenne River Sioux tribes. *Am J Public Health*. 2005;95:2129-2132.
54. Sarfaty M, Turner CH, Damotta E. Use of a patient assistant to facilitate medical visits for Latino patients with low health literacy. *J Commun Health*. 2005;30:299-307.
55. Seek AJ, Hogle WP. Modeling a better way: navigating the healthcare system for patients with lung cancer. *Clin J Oncol Nurs*. 2007;11:81-85.
56. Steinberg ML, Fremont A, Khan DC, et al. Lay patient navigator program implementation for equal access to cancer care and clinical trials: essential steps and initial challenges. *Cancer*. 2006;107:2669-2677.
57. Tingen MS, Weinrich SP, Heydt DD, Boyd MD, Weinrich MC. Perceived benefits: a predictor of participation in prostate cancer screening. *Cancer Nurs*. 1998;21:349-357.
58. Weinrich SP, Boyd MD, Weinrich M, Greene F, Reynolds WA Jr, Metlin C. Increasing prostate cancer screening in African American men with peer-educator and client-navigator interventions. *J Cancer Educ*. 1998;13:213-219.
59. Wilcox B. How do you make a difference in patients' lives through your role as a nurse navigator [interview]? *ONS Connect*. 2007;22:12.
60. Wujcik D. Navigator role shows promise in decreasing cancer death rates for all populations [comment]. *ONS Connect*. 2007;22:5.
61. Bruce S. Taking the wheel. *ONS Connect*. 2007;22:8-13.
62. Decker GM. Navigators: the champions of finding solutions to healthcare barriers. *ONS Connect*. 2007;22:29-29.
63. Oluwole SF, Ali AO, Adu A, et al. Impact of a cancer screening program on breast cancer stage at diagnosis in a medically underserved urban community. *J Am Coll Surg*. 2003; 196:180-188.
64. Till JE. Evaluation of support groups for women with breast cancer: importance of the navigator role [serial online]. *Health Qual Life Outcomes*. 2003;1:16.
65. Ferrante JM, Chen P-H, Kim S. The effect of patient navigation on time to diagnosis, anxiety, and satisfaction in urban minority women with abnormal mammograms: a randomized controlled trial. *J Urban Health*. 2008;85:114-124.
66. Fischer SM, Sauaia A, Kutner JS. Patient navigation: a culturally competent strategy to address disparities in palliative care. *J Palliat Med*. 2007;10:1023-1028.
67. Schwaderer KA, Itano JK. Bridging the healthcare divide with patient navigation: development of a research program to address disparities. *Clin J Oncol Nurs*. 2007;11:633-639.
68. Cancer Care Nova Scotia. Cancer Patient Navigation Evaluation: Final Report. Halifax, Nova Scotia, Canada: Cancer Care Nova Scotia; 2004.
69. Fowler T, Steakley C, Garcia AR, Kwok J, Bennett LM. Reducing disparities in the burden of cancer: the role of patient navigators [serial online]. *PLoS Med*. 2006;3: e193.
70. National Cancer Institute. Community update: patient navigator program guides underserved cancer patients [serial online]. *NCI Cancer Bull*. 2005;2:8.
71. Commission for Case Manager Certification. Case Management Practice. Available at: http://www.ccmcertification.org/pages/136body.html. Accessed March 13, 2007.
72. National Cancer Institute. Dictionary of Cancer Terms. Bethesda, Md: National Cancer Institute; 2007. Available at: http://www.cancer.gov/Templates/db_alpha.aspx?CdrID=44534. Accessed March 13, 2007.
73. Farber JM, Deschamps M, Cameron R. Investigation and Assessment of the Navigator Role in Meeting the Information, Decisional and Educational Needs of Women with Breast Cancer in Canada: Canadian Breast Cancer Initiative: Health Canada. Ottawa, Ontario: Minister of Public Works and Government Services Canada; 2002.
74. National Cancer Institute. Center to Reduce Cancer Health Disparities. Patient Navigator Research Program. Bethesda, Md: National Cancer Institute, 2007. Available at: http://crchd.cancer.gov/pnp/pnrp-index.html. Accessed December 31, 2007.
75. Calhoun EA, Whitley EM, Esparza A, et al. A National Patient Navigator Training Program. Presentation at the American Cancer Society Symposium, New Orleans, Louisiana, April 18-20, 2007.
76. Sheppard VB, Figueiredo M, Canar J, et al. Latina a Latina (SM): developing a breast cancer decision support intervention. *Psychooncology*. 2008;17:383-391.
77. Ward E, Halpern M, Schrag N, et al. Association of insurance with cancer care utilization and outcomes. *CA Cancer J Clin*. 2008;58:9-31.

Improving Bedside Teaching: Findings from a Focus Group Study of Learners

Keith N. Williams, MD, MSME, EdM, Subha Ramani, MBBS, MMEd, MPH, Bruce Fraser, PhD, and Jay D. Orlander, MD, MPH

Abstract

Purpose
Literature reviews indicate that the proportion of clinical educational time devoted to bedside teaching ranges from 8% to 19%. Previous studies regarding this paucity have not adequately examined the perspectives of learners. The authors explored learners' attitudes toward bedside teaching, perceptions of barriers, and strategies to increase its frequency and effectiveness, as well as whether learners' stages of training influenced their perspectives.

Method
Six focus group discussions with fourth-year medical students and first- or second-year internal medicine residents recruited from the Boston University School of Medicine and Residency Program in Internal Medicine were conducted between June 2004 and February 2005. Each 60- to 90-minute discussion was audiotaped, transcribed, and analyzed using qualitative methods.

Results
Learners believed that bedside teaching is valuable for learning essential clinical skills. They believed it is underutilized and described many barriers to its use: lack of respect for the patient; time constraints; learner autonomy; faculty attitude, knowledge, and skill; and overreliance on technology. Learners suggested a variety of strategies to mitigate barriers: orienting and including the patient; addressing time constraints through flexibility, selectivity, and integration with work; providing learners with reassurance, reinforcing their autonomy, and incorporating them into the teaching process; faculty development; and advocating evidence-based physical diagnosis. Students focused on the physical diagnosis aspects of bedside teaching, whereas views of residents reflected their multifaceted roles as learners, teachers, and managers.

Conclusions
Bedside teaching is valuable but underutilized. Including the patient, collaborating with learners, faculty development, and promoting a supportive institutional culture can redress several barriers to bedside teaching.

Acad Med. 2008; 83:257–264.

> There is a general uneasiness both in the minds of the public and also in the practicing physician, that future specialists in internal medicine will become mostly reliant on laboratory, computerized, nonpersonal techniques of management, and the patient as a human being with emotional and psychological aspects will be forgotten. If such physicians are to come into being, it must be due to the kind of training and environment to which they are exposed in their years in medical school.
>
> —L.H. Nahum

Bedside teaching, clinical teaching done in the presence of a patient, has been a fundamental component of medical training in the United States since the institution of modern methods of instruction in the late 19th century. Although few data exist regarding the effectiveness of bedside teaching, many medical educators espouse its value in training physicians.[1–17] It seems logical to assume that clinical skills related to physician–patient communication, physical examination, clinical reasoning, and professionalism are better learned at the bedside than in a classroom. Nevertheless, the proportion of clinical educational time devoted to bedside teaching has ranged from 8% to 19% since the 1960s.[18–23] Although medical educators have speculated on the reasons for this paucity, few studies have examined the issue systematically. Existing studies have had a limited scope or have explored bedside teaching only from the perspective of teachers.[24–28] Although numerous guidelines advocating specific bedside teaching strategies have been published, it is unclear whether the perspectives of learners influenced their development.[29–42]

Learners are likely to have unique and valuable perspectives on bedside teaching; any effort to increase or improve bedside teaching should consider their views. Our objectives were to explore learners' attitudes toward bedside teaching, perceptions of barriers, and strategies to increase its frequency and effectiveness. We included learners at different stages of training to assess whether experience influenced perspective.

Method

With the exception of the principal investigator (K.W.), all coinvestigators had qualitative research experience before the conduct of this study, and one (B.F.) taught a graduate course on qualitative research. We conducted six focus group discussions between June 2004 and February 2005 with students from the Boston University School of Medicine and residents from the Boston University Residency Program in Internal Medicine. All prospective participants received an e-mail letter of invitation. Participation was voluntary and confidential, and verbal consent was obtained from all participants. We audiotaped discussions and transcribed them verbatim. The institutional review board of the Boston University Medical Center approved the research protocol.

Please see the end of this article for information about the authors.

Correspondence should be addressed to Dr. Williams, c/o Dr. Ramani, Boston University School of Medicine, 715 Albany St., Vose 320, Boston, MA 02118; telephone: (617) 638-7985; fax: (617) 638-7905; e-mail: (Keith.Williams@bmc.org).

We sought varied perspectives by recruiting participants at different stages of training. Groups one and two consisted of fourth-year students. Groups three and five consisted of "first-year" internal medicine (IM) residents, and groups four and six of consisted of "second-year" IM residents; by definition, the former had completed one year of postgraduate study, and the latter had completed two.

We defined bedside teaching as clinical teaching in the presence of a patient. We constructed open-ended questions to explore learners' experiences and opinions regarding bedside teaching. We asked whether they learned from bedside teaching and, if so, what they had learned. We asked about the quality and quantity of bedside teaching they received. Finally, we solicited their views on barriers to bedside teaching and suggestions on strategies to increase its frequency and effectiveness. The interviewer (K.W.) pursued relevant themes and sought clarification or elaboration as required. Participants had ample opportunity to express unsolicited opinions.

Focus group discussions were 60 to 90 minutes in duration. The principal investigator (K.W.) conducted all interviews and, using standard qualitative methods, coded the transcripts.[43] We grouped coded passages into major categories and identified prominent themes that emerged. We also identified contrasting responses across the experience levels of participants.

Results

Thirty-three students and residents participated in one of six focus group discussions (Table 1). All residents had attended medical school in the United States or Canada. Coded passages generated several categories: value of bedside teaching, quantity and quality of bedside teaching, barriers to bedside teaching, and strategies to increase and improve bedside teaching. Differences between students' and residents' views were apparent and reflected differences in their roles. When applicable, we have noted the level of learner. Statements represent the views of learners and not the authors.

Value of bedside teaching

Learners believed bedside teaching is valuable, if not essential, for learning skills relating to physician–patient communication, physical examination, clinical reasoning, and professionalism (List 1). They reported that observing the resident or attending physician interact with patients is often instructive. Learners indicated that patients also benefit from bedside teaching encounters, and they emphasized that discussions of topics not directly related to patient care are more appropriately taught elsewhere.

> It's very powerful if you see the example on an actual person, and especially if you know more about their story, their background, you're more likely to take something away from that experience, whether it be some kernel of knowledge about a disease or a certain way of interacting with patients. (*Fourth-year medical student*)

List 1

Knowledge and skills learned with bedside teaching

- Obtaining a medical history
- Performing a physical examination
- Generating a differential diagnosis
- Formulating a management plan
- Applying clinical reasoning
- Communicating effectively
- Exhibiting professional bedside demeanor
- Demonstrating empathy
- Performing diagnostic and therapeutic procedures
- Acquiring knowledge about medical instrumentation

Quantity and quality of bedside teaching

Learners stated that bedside teaching is underutilized and that there are missed opportunities for bedside teaching. They reported that the quantity and quality of bedside teaching vary greatly among faculty and, for students, between clerkships. Students noted the most bedside teaching during their IM clerkships, but even there, quantity and quality vary.

> We do [bedside teaching] rarely because I feel like when it happens it stands out so much. (*Second-year IM resident*)

Barriers to bedside teaching

Viewing bedside teaching as the interplay of patient, teacher, and learner in the context of the learning environment, barriers were classified as personal, interpersonal, or environmental (Table 2).

Table 1

Characteristics of 33 Volunteer Participants in Six Focus Groups on Bedside Teaching, Boston University (BU) School of Medicine and BU Residency Program in Internal Medicine, 2004–2005

				Age of participants			Students' specialty plans			Residents' specialty plans		
Group	Type	Men	Women	20–25	26–30	31–35	Medicine	Surgery	Other	Yes	No	Maybe
1	Fourth-year student	3	3	5	1	0	1	2	3	NA	NA	NA
2	Fourth-year student	4	1	3	2	0	2	1	2	NA	NA	NA
3	First-year resident	5	2	0	5	2	NA	NA	NA	5	0	2
4	Second-year resident	3	2	0	2	3	NA	NA	NA	5	0	0
5	First-year resident	3	3	1	4	1	NA	NA	NA	5	0	1
6	Second-year resident	3	1	0	3	1	NA	NA	NA	2	2	0

Table 2

Barriers to Bedside Teaching, and Strategies to Increase and Improve Bedside Teaching, Compiled from Six Focus Groups on Bedside Teaching, Boston University (BU) School of Medicine and BU Residency Program in Internal Medicine, 2004–2005

Category	Barrier	Strategy
Personal	Low initiative for teaching	Increase teaching initiative with institutional incentives
	Low teacher/learner expectations for teaching	Increase teacher/learner expectations with: • Explicit teaching expectations for teachers • Explicit learning objectives for students and residents
	Inadequate bedside teaching skills	Develop teaching skills through faculty development and resident training initiatives • Create a supportive learning environment (e.g., admission by teachers of own limitations/errors) • Acknowledge learners' needs • Plan teaching in a flexible manner to accommodate work schedules • Selectively and efficiently integrate teaching with work • Set time limits when teaching
	Inadequate clinical knowledge and/or skills (faculty)	Improve clinical knowledge and/or skills through faculty development initiatives (e.g., advanced training in evidence-based physical diagnosis)
Interpersonal	Lack of patient cooperation	• Request permission from the patient • Orient the patient to the dual purpose of the bedside session (i.e., patient care and teaching) • Include the patient in discussions and answer questions • Inform the patient about his/her care (i.e., patient education)
	Learners' desire for autonomy in patient care/fear of a compromised relationship with the patient	• Respect the learner–patient relationship • Negotiate an appropriate level of autonomy with learners • Create a supportive learning environment • Share teaching responsibility with team members
	Learner/patient fear of embarrassment/humiliation	Learner • Create a supportive learning environment Patient • Request permission from and orient the patient • Include and inform the patient
Environmental	Lack of time attributable to high patient volume and turnover	• Reduce service caps on the number of patients admitted and/or managed • Create nonteaching services for patient overflow
	Competing responsibilities of faculty	Reduce or eliminate competing demands on faculty such as outpatient clinical duties and research responsibilities
	Deficient institutional expectations/incentives for teaching	Increase institutional expectations/incentives for teaching with: • Explicit teaching expectations/incentives for faculty/residents • Explicit learning objectives for residents/students
	Inadequate institutional recognition of teaching	Enhance institutional recognition of teaching with legitimate rewards for excellence in teaching
	Devaluation of clinical skills by technology	Emphasize evidence-based clinical diagnosis through faculty development and resident training initiatives
	Interruptions during rounds	No strategy offered
	Lack of privacy in multipatient room	No strategy offered
	Lack of space within patient room	No strategy offered
	Excessive noise	No strategy offered

Personal barriers are factors attributable to individuals, whereas interpersonal barriers represent aspects of the relationship between at least two individuals. Environmental barriers denote contextual factors that influence bedside teaching. The learning environment includes cultural aspects of the learning institution as well as structural and functional aspects of the patient-care environment. Several overarching themes emerged from the data: lack of respect for the patient; time constraints; learner autonomy; faculty attitude, knowledge, and skill; and overreliance on technology.

Lack of respect for the patient. Learners expressed concern for patients' welfare and recognized that their own education is secondary to patient care. Bedside discussion of sensitive issues, such as substance abuse, mental health, and sexuality-related topics, could embarrass the patient and result in the elicitation of

inaccurate information. Discussion of diagnostic possibilities could cause undue alarm, particularly conversations about fatal diseases such as cancer. Bedside deliberation of management plans could lead patients to lose confidence in the medical team if they witnessed disagreements. Learners attested that patients are rarely asked permission or oriented to bedside teaching. Most concerning to learners are situations in which the patient is marginalized during a bedside discussion, as manifested by a clinician's failure to seek patient input, explain medical terminology, or answer questions.

> I've seen attendings or residents exclude the patient when they're bedside teaching, and patients find that really offensive because it's their body, it's their story, and they're marginalized while they're being used for teaching, whereas if the patient's included then it's great for everybody and it's a really effective learning tool. (*First-year IM resident*)

Time constraints. Most learners believed that time constraints significantly limit bedside teaching. Contributing to this perception are a high resident workload and the observation that attendings often maintain significant ambulatory, administrative, or research responsibilities during their ward service, thus reducing time for learner interaction and bedside teaching. Learners preferred to avoid extended teaching encounters when overwhelmed by workload or fatigue.

> You duck out and you're not involved in the actual teaching at the bedside because you need the computer and the phone in the hallway to get things done so that things are happening earlier in the day and you're not discharging someone at six in the afternoon. (*First-year IM resident*)

Interestingly, some learners questioned why bedside teaching should take more time than that required for teaching in other settings. Such speculation led them to conclude that there is not so much a lack of time for bedside teaching, but for all teaching.

> Theoretically, bedside teaching shouldn't take any longer. What we're actually saying is there's not enough time for teaching in general, rather than for bedside teaching. (*First-year IM resident*)

Learner autonomy. Advanced residents expressed concern that bedside teaching compromises the relationship between learner and patient. They feared that bedside demonstration of deficiencies causes patients to lose confidence in the learner as clinician, or in the advanced resident as team leader. They also feared that their attendings would usurp their authority to manage the team. Although many learners voiced discomfort at admitting "I don't know" to their patients, some accepted the necessity, if not inevitability, of making such an admission. Some believed that lack of harmonious team dynamics contributes to these difficulties.

> It is uncomfortable when you're the resident and the patient knows that you're the one that's there all the time and somebody's going through this lengthy teaching episode with you, because I feel they want to be comforted by the fact that you know something, which you do, but there's that line where it can be uncomfortable. (*First-year IM resident*)

Ironically, some residents believed that attendings sometimes allow excessive autonomy. In such circumstances, all teaching suffers, not just that performed at the bedside.

> We crave autonomy, but it's a balance. I'm at the point now where I don't want as much autonomy, I actually want to interact more with my attendings. I had an attending this past month who gave me too much autonomy. It was fun, it was easier, but I didn't learn much. (*Second-year IM resident*)

Faculty attitudes, knowledge, and skill. Learners believed that all attendings have something to teach at the bedside. There was speculation that the reluctance of attendings to engage in bedside teaching derives more from lack of teaching skill than lack of clinical competence.

> I definitely ran into some people where I'd ask them questions and I could tell that I was making them uncomfortable, even though they had lots to offer me. That was what was really strange; they didn't realize that I'm asking at such a basic level that they have lots to teach me. (*Fourth-year medical student*)

Overreliance on technology. Given the ever-increasing technological options for diagnosis and treatment, some learners questioned the role and importance of proficiency in bedside diagnostic skills if further testing is done regardless of clinical impression. This loss of faith in traditional skills led to speculation about whether efforts to learn them are futile.

> So much of medicine now with radiology studies and blood tests is so algorithm driven that a lot of this other stuff doesn't matter anymore. I know what the plan's going to be before I see the patient. A lot of bedside teaching is gone away because of that. (*First-year IM resident*)

> One intern used to call the CT scanner the "doughnut of truth." It's kind of revealing. It's like saying that you don't need to lay your hands on the patient, just plop them down on the thing. (*Second-year IM resident*)

However, many learners believed that, in spite of technology, patients still expect dialogue with and examination by their physicians and are disappointed and indignant when clinical interactions are inadequate. Some learners desired international medical experiences where they could obtain traditional clinical teaching in the absence of the influence of technology.

> I don't think we'll ever get away from the physical exam. Even if you don't have to listen to their lungs and learn anything, the patients feel so much more of a connection to you when you do. It's more than just learning the physical exam. It's learning patient interaction, too. (*First-year IM resident*)

Strategies to increase and improve bedside teaching

Learners provided many insightful recommendations to increase and improve bedside teaching (Table 2). Strategies addressing the barrier themes noted previously are discussed below.

Orient and include the patient. Although patients were rarely described as uncooperative, learners emphasized the need to orient patients and request permission before teaching. Explaining the purpose of a bedside teaching encounter and requesting permission to observe or examine were identified as important signs of respect that foster trust and cooperation.

> The most important thing is being able to develop a rapport with your patients, and making them feel like you're not forcing something on them, where it's a comfortable environment. If an attending or resident has the ability to establish a relationship with the patient that's reasonable, you can get a lot more out of that situation because the patient is more willing to participate and the students will feel more comfortable in that setting, too. (*Fourth-year medical student*)

Learners believed that inclusion of patients permits clarification of historical facts and validation of exam findings. It allows the medical team to educate patients about their conditions and options for care. They felt that establishing rapport with patients could alleviate concerns regarding discussion of sensitive topics and prevent misunderstandings or alarm with regard to discussion of diagnostic possibilities or management plans. Patients might also have the satisfaction of contributing to the education of future physicians.

> The attending did a good job of making the patient feel we cared about her. He set a good example of making her feel like a whole person and not a specimen, in the end reassuring her about her findings: "This isn't new, this is something that we've already known you have, and what I'm talking about here isn't anything for you to worry about, and it's consistent with the diagnosis you have and nothing's changed." I thought that was a good closure to the teaching. (*Fourth-year medical student*)

Address time constraints through flexibility, selectivity, and integration. Although many time-related factors are beyond their control, learners provided several suggestions to address the effect of time constraints. Faculty accommodation of the call schedule and resident workload was considered paramount. Learners stated that it is senseless to engage in prolonged postcall teaching rounds when the team is too fatigued and distracted to engage in a meaningful learning experience. Faculty teaching at the bedside should be selective and limited in duration. Paradoxically, some residents suggested that faculty participate regularly in work rounds; they recognized the efficiency of combining work and learning, especially when it obviates the need to round later with the attending. Residents advocated a decrease in the number of patients they manage on the teaching ward. Considering the amount of clerical work associated with each patient, even slight reductions in census numbers would increase the time available for teaching.

> [It] is very important to prioritize, to pick one or two pearls on the patients that you're interested in and emphasize that, because we're not going to remember more than that, and we usually don't have time for more than that. (*First-year IM resident*)

Learners emphasized the value of having attendings available on the ward on a consistent basis. Attending time on the ward should be "protected" from competing responsibilities.

> It really helps when the focus of that person's day is to take care of patients and to do teaching . . . it makes a big difference . . . rather than someone who's got their hands in so many things, their mind might be in as many places. (*Fourth-year medical student*)

Provide learners with reassurance, reinforce their autonomy, and incorporate them into the teaching process. To alleviate their anxiety at the bedside, learners believed that simple reassurance by the attending physician is often sufficient. They advocated the establishment of a positive learning environment in which acknowledgement of deficiencies and errors is accepted as an inevitable, if not essential, aspect of the learning process. In this context, bedside questioning is seen as a Socratic exercise in learning, not as "pimping" with an intent to demoralize or cause embarrassment. Some believed that harmonious team dynamics facilitate acceptance of one's own limitations.

> If teachers can set the tone and what the expectations are and say that "it is okay to make mistakes, we all make mistakes, but the great doctors are the ones who take those mistakes and use them to improve themselves," that's the best way to learn in that stressful environment. (*Second-year IM resident*)

Some residents suggested that autonomy, although important for professional growth, could sometimes be counterproductive in its effect on opportunities for learning, because autonomy often correlates inversely with attending participation. They recognized the need to balance their roles as managers and learners, although it is often difficult to negotiate the appropriate level of autonomy with attendings. They believed that distributing teaching responsibility to all team members and creating a collaborative learning environment could minimize compromise of the professional relationship between learner and patient.

> One solution is to equalize the teaching on the team. Just because you're ahead of someone else, like the attending's above you or you're above the intern, doesn't mean you're going to know more about every topic than the medical student. (*First-year IM resident*)

Develop faculty attitudes, knowledge, and skill for bedside teaching. Learners suggested faculty and resident training to develop effective bedside teaching skills. Such training could alleviate the anxiety related to bedside teaching content, such as physical examination skills. They advocated a variety of bedside teaching strategies, including assessment of learners' needs, role modeling, selective and explicit instruction, learner evaluation and feedback, and distribution of teaching responsibility. Legitimate institutional incentives for proficient teaching should be available.

> Everything counts the minute you walk into the patient's room. Everything you do is being watched, whether it's something you say, or it's the way you approach the patient, the way you sit by the bed, or just the way you're ignoring what the patient's saying. Teachers should be very sensitive to that issue alone. It's not just the verbal aspect of teaching. (*Second-year IM resident*)

Advocate evidence-based physical diagnosis. Learners believed that the indifference of the medical establishment towards physical diagnosis skills derives from lack of emphasis in training, and they suggested that faculty and resident training initiatives could improve these skills. They encouraged participation in international medical experiences in which technology does not play a central role in the diagnosis and treatment of illness.

> Some things in physical exam are actually useful. There's some literature on the prognosis implied in a certain physical exam finding. Us[e] that to say, "these things are important and it can actually guide the management." (*Second-year IM resident*)

Table 2 lists these and additional strategies to increase and improve bedside teaching.

Contrasting student and resident perspectives

Residents' beliefs, such as the desire for autonomy in patient care and for a collaborative learning environment, were more pragmatic than those of students. This pragmatism stemmed from two major differences between students and residents. First, residents viewed bedside teaching, and teaching in general, from

the vantage points of both learner and teacher. Second, their views were influenced by work responsibilities and a desire to have a reasonable quality of life, even during training. Teaching initiatives that fail to recognize these differences are often unsuccessful.

Students focused on the physical diagnosis aspects of bedside teaching to a greater degree than did residents. First-year residents were overwhelmed with the responsibilities of daily work, and thus found bedside teaching, and perhaps all teaching, to be another demand on their limited time, and they reported that they often felt too distracted to learn. In the second year of residency and beyond, residents recognized that their role as a team leader allowed them opportunities to influence the frequency and form of bedside teaching rounds.

> There were definitely times where somebody said the word[s] "attending rounds" and I was ready to shoot myself. I was completely disinterested and in fact angry that that was what somebody wanted to do when I had a million other things, and it was only going to keep me in the hospital really late. Some people just don't have any understanding of what's going on around them. (*First-year IM resident*)

> The resident's attitude permeates the team, so you can create a local environment of eagerness and motivation to learn. (*Second-year IM resident*)

Discussion

Our learners confirmed faculty beliefs that bedside teaching is valuable for learning essential clinical skills, such as those related to physician–patient communication, physical examination, clinical reasoning, and professionalism.[27,28] In the absence of studies validating the effectiveness of bedside teaching, this affirmation by learners is important. Their recognition that they learn by observing more experienced clinicians interact with patients supports the use of role modeling, an implicit form of teaching, at the bedside. Although they had concerns for the patient and their own psychological well-being during the bedside teaching encounter, they identified strategies to avoid potential harm. We are encouraged, given the learners' beliefs that patient inclusion and faculty development could rapidly enhance the frequency and effectiveness of bedside teaching for the benefit of all.

One of our most striking findings is the recognition by learners that, for the patients and themselves, sensitivity in the interpersonal aspects of bedside teaching is paramount. A poorly executed bedside teaching encounter disrespects patients and compromises learners in their roles as clinicians and managers, thus diminishing their perceived autonomy. Fear of the consequences of poor interpersonal communication during bedside teaching is prevalent. Simple strategies to avoid these pitfalls, such as orienting patients to the process, and explicit acknowledgment of human limitations, can be easily incorporated by faculty.

Although lack of time is frequently described as a major barrier to bedside teaching, some learners believed that this is more perception than reality. In fact, the data suggest that when time is limited, all teaching is compromised, not just bedside teaching. More importantly, learners suggested that properly executed bedside teaching could be integrated within typical clinical activities, such as work rounds, allowing efficient time use.

Autonomy was a major concern for residents. They avoid teaching situations that might jeopardize their role as manager and the students' or first-year residents' role as caregiver. A collaborative approach to teaching helps to preserve the integrity of their semiautonomous roles. Learners found it difficult to admit "I don't know" during a bedside teaching encounter. Reassurance by attendings and the establishment of a positive learning environment and harmonious team dynamics can alleviate these concerns.

Learners speculated that faculty might be reluctant to teach at the bedside because of a lack of teaching skills rather than a lack of clinical competence. In their opinion, the belief that technology has supplanted the medical history and physical examination undermines bedside teaching. Faculty development could address both of these issues.

The differences in the perspectives of students and residents regarding bedside teaching are provocative. One could predict both the tremendous appeal of bedside teaching to students as they learn the skills of clinical medicine and the pragmatism of first-year residents inundated with work. However, second-year residents' expectation for a more collaborative approach to teaching is somewhat unexpected and compelling. This expectation stemmed from a need for autonomy as they embraced their newfound leadership role. Their desire for a collaborative approach has profound implications: if given the opportunity to influence the timing, content, and process of bedside teaching, residents could have a key role in promoting the regular occurrence of such teaching.

Our findings complement previous literature reports regarding learners' perspectives on bedside teaching. In a study by Nair et al,[26] learners were found to believe that bedside teaching is a "valuable way to develop professional skills." More than 90% of the learners believed that bedside teaching is effective for learning communication, history-taking, and physical examination skills. Between 41% and 65% stated that they do not receive sufficient bedside teaching. Our findings also complement the views of teachers. In another study by Nair et al,[27] 95% of teachers agreed that bedside teaching is an effective way to develop professional skills, and more than 80% believed that it is effective for learning communication, history-taking, and physical examination skills. Comparison with a study by Ramani et al[28] reveals a striking similarity between the perspectives of teachers and the views of our learners; teachers concurred with the five overarching themes we describe in this paper, including learner autonomy, as suggested by a "fear of undermining housestaff." They also suggested strategies to increase and improve bedside teaching that are remarkably consistent with those of our learners, such as orienting the patient, establishing a positive learning environment, and treating the learner as primary caregiver for the patient. The suggestions of our learners are compatible with the bedside teaching recommendations of various educators.[29–42] The "model of best bedside teaching practices" by Janicik and Fletcher,[42] which describes three domains of effective bedside teaching skills (attending to patient comfort, focused teaching, and group dynamics), addresses several of the important findings in our study.

Bedside teaching, rather than being an antiquated mode of clinical instruction

from a pretechnological era, is consistent with modern education theory. Specifically, it is consistent with the experiential learning principles of the progressive movement that began in medical education during the late 19th century, as well as with the principles of adult learning as defined by Malcolm Knowles.[44,45] Most striking, however, is the compatibility of bedside teaching with the modern theory of situated cognition, or contextual learning, which states that the learning of knowledge is inherently dependent on the context in which it is learned; that is, "knowledge is situated, being in part a product of the activity, context, and culture in which it is developed and used."[46] William Osler[4] was aware of the advantages of contextual learning: "In what may be called the natural method of teaching, the student begins with the patient, continues with the patient, and ends his studies with the patient, using books and lectures as tools, as means to an end." Our learners were aware of this "contextual" advantage, as demonstrated by their views regarding the value of bedside teaching. One may logically ask, how else is a student of medicine to become a professional, let alone learn the appropriate manner of speaking with, touching, and comforting a patient, if not in the context of the bedside teaching encounter?

This study has several limitations. We recruited students from only one school of medicine, and residents from only one specialty at a single academic medical center. Our study explored learners' opinions about bedside teaching but did not address whether increasing or improving bedside teaching would lead to better learning outcomes or enhanced patient care. Although there is always the potential for interviewer bias to influence the views of participants during focus group discussions, we attempted to reduce this possibility by limiting the interviewer's dialogue to questions and clarifications and by avoiding expression of opinion. We also avoided asking whether learners "liked" bedside teaching; we focused instead on questions concerning value, barriers, and strategies.

Future research should determine whether faculty development directed at improving bedside clinical and teaching skills could enhance bedside teaching and should focus on a variety of learner outcomes, including knowledge retention, skill proficiency, and professionalism. Meanwhile, the results of this study support ongoing efforts to promote faculty development in teaching and to facilitate an institutional culture and environment conducive to the regular occurrence of bedside teaching.

We believe clinical education that incorporates substantial bedside teaching is an effective approach to fulfilling the public interest of training intelligent, skilled, and compassionate clinicians. Including the patient, collaborating with learners, developing faculty skills, and promoting a supportive institutional culture can redress a variety of barriers to bedside teaching. In the end, "no books, no tapes, no audio-visual aids, no seminars, no avant-garde philosophy will ever be substitutes for the discipline of the bedside medicine—the one-to-one situation where tradition, humanity, art and science are blended."[6]

Dr. Williams is clinical instructor of medicine, Boston University School of Medicine, Boston, Massachusetts. When this research was conducted, he was a general internal medicine clinician–educator fellow, Boston University School of Medicine MET Program, Boston, Massachusetts.

Dr. Ramani is associate professor of medicine and director of clinical skills development, Internal Medicine Residency Program, Boston University School of Medicine, Boston, Massachusetts.

Dr. Fraser is professor of linguistics and education, Boston University Applied Linguistics Program and School of Education, Boston, Massachusetts.

Dr. Orlander is associate professor of medicine, Boston University School of Medicine, and associate chief medical service, VA Boston Health Care System, Boston, Massachusetts.

Acknowledgments

This research was supported in part by the HRSA Faculty Development Training Grant # D55HP00215.

Disclaimer

Preliminary results of this research were presented orally at the 2005 New England SGIM Regional Meeting, Hanover, New Hampshire, and by poster at the 2005 SGIM Annual Meeting, New Orleans, Louisiana, and the 2005 AAMC Annual Meeting, Boston, Massachusetts.

References

1 Nahum LH. Medical teaching rounds. Conn Med. 1966;30:225–226.
2 Hill GL, Bowen JC, Copeland EM, MacFadyen BV, Duke JH Jr, Dudrick SJ. Teaching in clinical surgery: At the bedside or in the seminar room? J Med Educ. 1977;52:595–597.
3 Hill DA, Lord RSA. Complementary value of traditional bedside teaching and structured clinical teaching in introductory surgical studies. Med Educ. 1991;25:471–474.
4 Osler W. The natural method of teaching the subject of medicine. JAMA. 1901;36:1673–1679.
5 Futcher TB. The importance of bed-side study and teaching. Can Med Assoc J. 1935;32:357–364.
6 Bedside medicine—the Oslerian reliquary. Med J Aust. 1971;1:1251–1252.
7 Linfors EW, Neelon FA. The case for bedside rounds. N Engl J Med. 1980;303:1230–1233.
8 Morgan WL Jr. Bedside teaching. Trans Am Clin Climatol Assoc. 1981;93:164–171.
9 Mazzaferri EL. Bedside teaching in the preparation of physicians for the 21st century. Arch Intern Med. 1986;146:1912.
10 St. Geme JW Jr. Can we rediscover the bedside? Pediatrics. 1987;79:287–288.
11 Belkin BM, Neelon FA. The art of observation: William Osler and the method of Zadig. Ann Intern Med. 1992;116:863–836.
12 Fitzgerald FT. Bedside teaching. West J Med. 1993;158:418–420.
13 LaCombe MA. On bedside teaching. Ann Intern Med. 1997;126:217–220.
14 Thibault GE. Bedside rounds revisited. N Engl J Med. 1997;336:1174–1175.
15 Kroenke K, Omori DM, Landry FJ, Lucey CR. Bedside teaching. South Med J. 1997;90:1069–1074.
16 Langlois JP, Thach S. Teaching at the bedside. Fam Med. 2000;32:528–530.
17 El-Bagir M, Ahmed K. What is happening to bedside clinical teaching? Med Educ. 2002;36:1185–1188.
18 Payson HE, Barchas JD. A time study of medical teaching rounds. N Engl J Med. 1966;273:1468–1471.
19 Collins GF, Cassie JM, Daggett CJ. The role of the attending physician in clinical training. J Med Educ. 1978;53:429–431.
20 Tremonti LP, Biddle WB. Teaching behaviors of residents and faculty members. J Med Educ. 1982;57:854–859.
21 Shankel SW, Mazzaferri EL. Teaching the resident in internal medicine. JAMA. 1986;256:725–729.
22 Miller M, Johnson B, Greene HL, Baier M, Nowlin S. An observational study of attending rounds. J Gen Intern Med. 1992;7:646–648.
23 Elliot DL, Hickam DH. Attending rounds on in-patient units: Differences between medical and non-medical services. Med Educ. 1993;27:503–508.
24 Wang-Cheng RM, Barnas GP, Sigmann P, Riendl PA, Young MJ. Bedside case presentations: Why patients like them but learners don't. J Gen Intern Med. 1989;4:284–287.
25 Kroenke K, Simmons JO, Copley JB, Smith C. Attending rounds: A survey of physician attitudes. J Gen Intern Med. 1990;5:229–233.
26 Nair BR, Coughlan JL, Hensley MF. Student and patient perspectives on bedside teaching. Med Educ. 1997;31:341–346.
27 Nair BR, Coughlan JL, Hensley MF. Impediments to bed-side teaching. Med Educ. 1998;32:159–162.
28 Ramani S, Orlander JD, Strunin L, Barber TW. Whither bedside teaching? A focus

group study of clinical teachers. Acad Med. 2003;78:384–390.
29 McLeod PJ. A successful formula for ward rounds. Can Med Assoc J. 1986;134:902–904.
30 Kroenke K. Attending rounds: Guidelines for teaching on the wards. J Gen Intern Med. 1992;7:68–75.
31 Cox K. Planning bedside teaching—1. Overview. Med J Aust. 1993;158:280–282.
32 Cox K. Planning bedside teaching—2. Preparation before entering the wards. Med J Aust. 1993;158:355–357.
33 Cox K. Planning bedside teaching—3. Briefing before seeing the patient. Med J Aust. 1993;158:417–418.
34 Cox K. Planning bedside teaching—4. Teaching around the patient. Med J Aust. 1993;158:493–495.
35 Cox K. Planning bedside teaching—5. Debriefing after clinical interaction. Med J Aust. 1993;158:571–572.
36 Cox K. Planning bedside teaching—6. Reflection on the clinical experience. Med J Aust. 1993;158:607–608.
37 Cox K. Planning bedside teaching—7. Explication of the clinical experience. Med J Aust. 1993;158:789–790.
38 Cox K. Planning bedside teaching—8. Deriving working rules for next time. Med J Aust. 1993;159:64–65.
39 Ende J. What if Osler were one of us? Inpatient teaching today. J Gen Intern Med. 1997;12(suppl 2):41–48.
40 Mooradian NL, Caruso JW, Kane GC. Increasing the time faculty spend at the bedside during teaching rounds. Acad Med. 2001;76:200.
41 Ramani S. Twelve tips to improve bedside teaching. Med Teach. 2003;25:112–115.
42 Janicik RW, Fletcher KE. Teaching at the Bedside: A New Model. Med Teach. 2003;25:127–130.
43 Patton MQ. Qualitative research and evaluation methods. Thousand Oaks, Calif: Sage Publications; 2002.
44 Ludmerer KM. Learning to Heal. The Development of American Medical Education. Baltimore, Md: The Johns Hopkins University Press; 1985.
45 Knowles MS. Self-Directed Learning: A Guide for Learners and Teachers. New York, NY: Association Press; 1975.
46 Brown JS, Collins A, Duguid P. Situated cognition and the culture of learning. Educ Res. 1989;18:32–42.

Correction

In the article "Beyond the Dual Degree: Development of a Five-Year Program in Leadership for Medical Undergraduates," in the January 2008 issue of *Academic Medicine*, an author's name was misspelled. The correct authors are: Gerald E. Crites, MD, MEd, James R. Ebert, MD, MBA, and Richard J. Schuster, MD, MMM.

CLINICAL RESEARCH STUDY

THE AMERICAN JOURNAL of MEDICINE®

Secular Trends in Alcohol Consumption over 50 Years: The Framingham Study

Yuqing Zhang, DSc,[a,b] Xinxin Guo, MPH,[a] Richard Saitz, MD, MPH,[c,d] Daniel Levy, MD, MPH,[a,e] Emily Sartini, MA,[e] Jingbo Niu, DSc,[b] R. Curtis Ellison, MD[a]

[a]Section of Preventive Medicine and Epidemiology, Evans Department of Medicine, [b]Clinical Epidemiology Research and Training Unit, [c]Clinical Addiction Research and Education Unit, Section of General Internal Medicine, Department of Medicine, Boston University School of Medicine and Boston Medical Center, Mass; [d]Department of Epidemiology and Youth Alcohol Prevention Center, Boston University School of Public Health, Mass; [e]NHLBI Framingham Heart Study, Framingham, Mass.

ABSTRACT

BACKGROUND: Population trends in patterns of alcohol use are important data for policymakers but are generally based on repeated cross-sectional surveys.

METHODS: We used self-reported alcohol consumption data collected repeatedly over 50 years (1948-2003) among 8600 Framingham Heart Study participants to determine patterns of alcohol use and disorders according to sex, age, and birth cohorts.

RESULTS: Among drinkers, there was a decrease across succeeding birth cohorts in average alcohol intake: among individuals between ages 30 and 59 years, age-adjusted mean intake was 30.6, 25.5, and 21.0 g/day for those born in 1900-1919, 1920-1939, and 1940-1959, respectively, in men ($P < .001$), and 14.2, 12.3, and 10.4 g/day, respectively, in women ($P < .001$). In all birth cohorts, proportion of abstinence increased and average consumption among drinkers decreased with age. Furthermore, proportion of moderate use was higher but heavy use was lower in the younger birth cohorts than in the older cohorts. The proportion of alcohol from beer decreased and that from wine increased with age for all cohorts. Among the 2 earlier birth cohorts, the cumulative incidence of an alcohol use disorder from age 40 to 79 years was much higher in men (12.8%) than in women (3.8%); it tended to be slightly higher among subjects born after 1920 than among those born 1900-1919.

CONCLUSIONS: We found a decrease in average intake and more wine consumption over the more than 50 years of follow-up. The cumulative incidence of alcohol use disorders, however, did not show a decrease.
 • *The American Journal of Medicine (2008) 121, 695-701*

KEYWORDS: Alcohol drinking; Alcohol-related disorders; Cohort studies; Drinking behavior; Epidemiology

The adverse effects on health and on society of excessive alcohol use have been well described.[1-5] On the other hand, numerous studies have shown that "moderate" alcohol consumption without heavy drinking episodes is associated with a reduced risk of most cardiovascular diseases[6-8] and many other diseases associated with ageing,[9-11]and with improved morbidity and mortality among the elderly.[12] Further, there is appreciation of the importance of the pattern of alcohol consumption in determining the net effects of alcohol consumption, with the greatest health benefits and fewest adverse effects being associated with regular moderate consumption without binge drinking.[13,14]

A number of studies have described trends in alcohol use, drinking patterns, and frequency of unhealthy use in the United States.[15,16] Studies have generally used data collected from sequential cross-sectional surveys, hospital discharge records, or sales and traffic data. The results suggest that the total amount of alcohol consumed in the United States has decreased slightly over recent decades, and that certain indicators of unhealthy use, such as death from driving while intoxicated or alcohol-related cirrhosis, have generally decreased since the middle of the 20th century,

These analyses were supported by a grant from the National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health (R01 AA013304). The Framingham Heart Study is funded through NHLBI Contract N01-HC-25195.

Requests for reprints should be addressed to Yuqing Zhang, DSc, Clinical Epidemiology Research and Training Unit, Boston University School of Medicine, 715 Albany Street, Room A203, Boston, MA 02118.

E-mail address: yuqing@bu.edu

0002-9343/$ -see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.amjmed.2008.03.013

although there have been recent increases in the prevalence of heavy episodic drinking among young people and certain adverse consequences.[17-19] To date, no study has prospectively examined secular trends for the amount of alcohol consumed and patterns of drinking among subjects in a well-defined population-based sample followed over many decades.

The Framingham Study has been monitoring alcohol use since recruitment of the original cohort of subjects began in 1948 and an introduction of an offspring cohort in 1971. In addition to repeated measures of amount and frequency of alcohol consumed recorded in the study database, a new endeavor, the Lifetime Health Study, is currently conducting a page-by-page review of all paper records available on each Framingham Study participant to better elucidate the pattern of drinking, seek evidence of consequences of alcohol use, and identify individuals with alcohol use disorders.

CLINICAL SIGNIFICANCE

- Over 50 years of follow-up, there was a decrease across succeeding birth cohorts in average alcohol intake, with more moderate and less heavy consumption.
- The proportion of alcohol from beer decreased and that from wine increased in succeeding cohorts and with age.
- Despite more favorable secular patterns of drinking, risk of alcohol dependence did not show a decrease, warranting continued efforts at its prevention and treatment.

The present analysis of 8600 participants in the Framingham Study whose record reviews have been completed seeks to determine the effects of sex, age, and birth cohort over time on the total amount of alcohol consumption among drinkers; the proportion of alcohol from specific beverages (beer, wine, and liquor); the prevalence of abstainers and, among drinkers, the proportion who were "moderate" drinkers and "heavy" drinkers; and the cumulative incidence of an alcohol use disorder from age 40 to age 79 years.

METHODS

The Framingham Heart Study began in 1948 in Framingham, Massachusetts.[20] The original cohort, hereafter referred to as the "Original Cohort," included 5209 subjects, aged 28-62 years at the first examination. Starting with the second examination, surviving participants have been examined biannually. At each examination, participants received a medical history interview, a physical examination, and a series of laboratory tests. In 1971-1974, examinations were offered for the children of the Original Cohort and their spouses, and a total of 5124 subjects were examined as part of the Framingham Offspring Study. The second examination in the Offspring Cohort occurred approximately 8 years after the baseline examination; subjects have been followed in 4-year cycles since then, with evaluations similar to those of the Original Cohort.

Assessment of Alcohol Consumption

Average Alcohol Consumption. Information on amount of alcohol consumption has been collected repeatedly from both the Original Cohort and the Offspring Cohort. At early examinations (up to the 7th examination) of the Original Cohort, subjects were asked how many 2-oz cocktails, 8-oz glasses of beer, and 4-oz glasses of wine they consumed in a month. At subsequent examinations (the 12th-15th, the 17th-23rd, and the 26th-27th examinations) of the Original Cohort and at all examinations (the 1st-7th examinations) of the Offspring Cohort, subjects were asked about the number of 1.5-oz cocktails, 12-oz glasses (or cans) of beer, and 4-oz glasses of wine they consumed in a week. Total alcohol consumption (grams per day) has been computed by multiplying the average amounts of alcohol in beer, wine, and mixed drinks times the amount drunk. We adjusted for secular changes that occurred in the late 1960s in the alcohol content of liquor commonly consumed (from 100% to 80% proof) and the type of wine generally consumed (from fortified to table wine), as well as a change in the average serving sizes of drinks to calculate the total ethanol content according to when the data were collected.

Heavy Episodic Drinking. In the later examinations of the Original Cohort (the 15th and 17th-23rd examinations) and starting at the 2nd examination of the Offspring Cohort, subjects were asked by trained interviewers prompted by an examination form to specify, "On average, what is your limit for number of drinks at one period of time?" Using this information, subjects reporting more than 4 drinks per occasion for men and more than 3 drinks per occasion for women were identified as having heavy episodic drinking.

Alcohol Use Disorders. Trained research assistants reviewed all available medical records, including Framingham Study questionnaires, hospital records, letters from private doctors, and other original source documents, for each study participant. An ICD-9 (International Classification of Diseases, 9th Revision) code was assigned for every disease or condition that could be coded. For the current analysis, alcohol use disorders included ICD-9 codes of 305.0 (alcohol abuse), 303.9 (alcohol dependence), and other diagnoses considered to be due exclusively to alcohol ("100% attributable"), such as alcohol withdrawal symptoms (291.81), delirium tremens (291.0), alcoholic cardiomyopathy (425.5), alcoholic cirrhosis (571.2), and alcohol detoxification therapy (94.62).

We also identified an additional group that did not meet criteria for ICD-9 coding for an alcohol use disorder, but review of their medical records identified words/phrases indicating a "possible" alcohol use disorder. The record

review leading to this designation included a search for key words such as "lost job because of drinking," "attends AA," "arrested for drunk driving," "advised to decrease my drinking," "physically injured because of drinking," and many others. To check the quality of record review, we randomly selected 450 records from the Original Cohort and 450 records from the Offspring Cohort and had them re-reviewed independently by both the project manager (ES) and the principal investigator (RCE). The proportion of agreement on key words indicating a possible disorder was 99% for both the Offspring Cohort and the Original Cohort.

Statistical Analysis

Subjects were grouped into 4 birth cohorts: born before 1900, and born in 1900-1919, 1920-1939, and 1940-1959. Because data on alcohol consumption were collected biannually in the Original Cohort and every 4 years in the Offspring Cohort, we divided age into 2-year age categories, with the first age category being 18-21 years and the last being ≥86 years of age. Data included in these analyses were those collected from 1948 through November 19, 2003 (end of the 27th examination) for the Original Cohort and from 1971 through October 26, 2001 (end of the 7th examination) for the Offspring Cohort.

We calculated the average amount of alcohol consumption among the drinkers and plotted it against age for each birth cohort for men and women separately. For comparison of birth cohorts that had overlapping data, we calculated age-adjusted average amount of alcohol consumption between ages 30 and 59 years for the 3 younger birth cohorts and between ages 60 and 79 years for the 3 older birth cohorts. We assessed the effect of age and birth cohorts on average total alcohol consumption using generalized estimating equations.[21]

We estimated the proportion of the total amount of alcohol consumed from beer, wine, and liquor by dividing the amount of beer, wine, and liquor consumed by subjects in a particular age group of the specific birth cohort by the total amount of alcohol consumed at the time by the same subjects.

We defined a participant as a nondrinker (abstainer) at a particular examination if he/she did not report any alcohol beverage intake at that examination. Among drinkers, we defined an individual as a "moderate" drinker if the subject's average amount of alcohol met current guidelines of the United States Department of Agriculture/Health and Human Services (an average of ≤24 g of alcohol per day for men and ≤12 g of alcohol per day for women) and no heavy episodic drinking (defined as >4 standard drinks per occasion for men and >3 drinks per occasion for women). Because data on maximum drinking on an occasion were collected only in more recent examinations of the Original Cohort, the analysis evaluating secular trends of "moderate" and "heavy" alcohol consumption among alcohol consumers was limited to the 3 younger birth cohorts. For each birth cohort, we obtained the age- and sex-specific prevalence of abstainers and the proportion of consumers deemed "moderate" and "heavy" drinkers, and examined their association with age and birth cohort using generalized estimating equations.

Finally, we estimated the cumulative incidence of an alcohol use disorder from age 40 to 79 years for the older cohorts (those with data extending to age 80 years). Because it is likely that very few subjects would become newly dependent after age 80 years, this is the closest we could come to giving the "lifetime risk" of an alcohol disorder for adults at age 40 years or older. Because some members of the Framingham Study were not recruited into the study until they were older than 40 years, our analyses could not include everyone in the current study who may have developed an alcohol use disorder at an earlier age. For the present analyses, we excluded any individual who was identified as having such a disorder before age 40 years. We also calculated the cumulative incidence by age 80 years of an alcohol use disorder when defined as having *either* an ICD-9 code for such a disorder *or* a key word indicating a possible alcohol use disorder.

RESULTS

The characteristics of the subjects at the time of their baseline examination (1948-1953 for the Original Cohort or 1971-1975 for the Offspring Cohort) are shown in Table 1. About one half of subjects were 40 years or younger at the time of their baseline examinations and over 40% were born before 1920. Approximately one quarter of participants did not graduate from high school. Prevalence of smoking was much higher in men than in women.

Table 1 Characteristics of Participants at Their Baseline Examination in The Framingham Heart Study, Framingham, MA, 1948-1975

Characteristics at Baseline	Men (n = 4097)	Women (n = 4541)
Age, years (%)		
30	18.7	20.6
31-40	32.7	32.7
41-50	30.3	29.9
>50	18.4	16.9
Birth cohort (%)		
<1900	12.2	13.0
1900-1919	30.1	30.9
1920-1939	34.5	30.9
1940-1959	23.2	25.1
Education (%)		
<High school	26.2	24.7
High school graduate	28.8	34.2
Some college	16.0	19.0
College graduate	29.0	22.1
Smoking (%)		
Never smoker	23.5	46.7
Past smoker not now	18.1	10.0
Current smoker	58.4	43.3

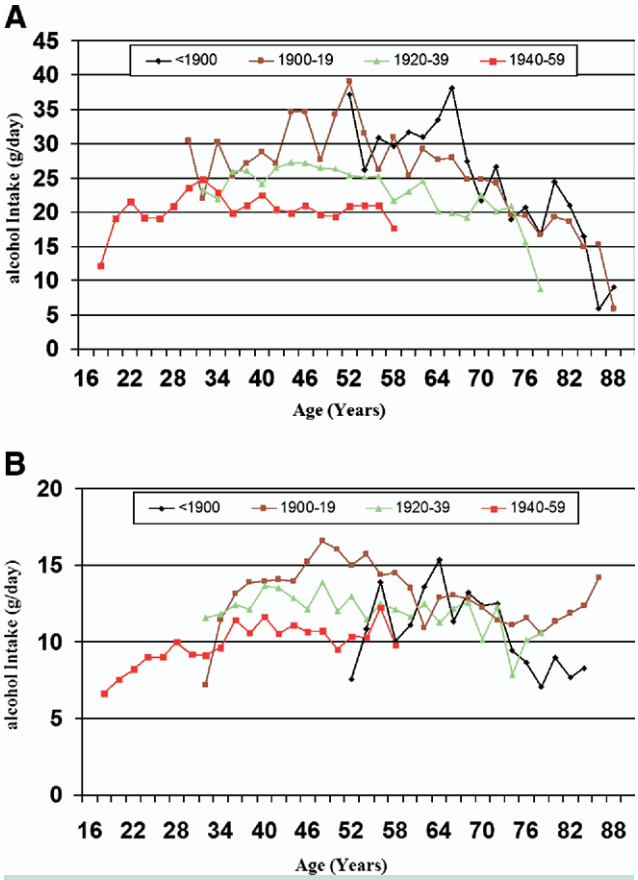


Figure 1 **(A)** Mean alcohol intake (grams/day) among 4097 male participants in the Framingham Study who consumed any alcohol, by age and birth cohort. **(B)** Mean alcohol intake (grams/day) among 4541 female participants in the Framingham Study who consumed any alcohol, by age and birth cohort.

The average amount of alcohol intake, in grams of alcohol per day, for individuals reporting any drinking is presented in Figure 1. In both men (Figure 1A) and women (Figure 1B), average consumption was lower among those in the younger birth cohorts compared with the older ones. Among men between ages 30 and 59 years, the age-adjusted mean alcohol consumption was 30.6, 25.5, and 21.0 g/day for those born in 1900-19, 1920-39, and 1940-59, respectively ($P < .001$). Corresponding figures for women were 14.2, 12.3, and 10.4 g/day, respectively ($P < .001$). Among subjects for whom data between ages 60 and 79 years were available, that is, those born before 1900, from 1900-1919, and 1920-1939, a similar pattern also was observed. Among these older men, age-adjusted mean alcohol consumption was 27.9 g/day for those born before 1900, 24.0 g/day for those born in 1900-1919, and 19.3 g/day for those born in 1920-1939 ($P < .001$). The corresponding amount of alcohol consumption for these older women changed less over time (11.6, 12.0, and 11.0 g/day, respectively, $P = .36$).

Among men reporting any alcohol intake at an examination, the proportion derived from beer showed a striking decrease with age for all birth cohorts, being 50% or more of their total alcohol for subjects until they reached their mid-30s and decreasing to approximately 25% by their mid-70s. For the youngest birth cohort (born 1940-1959), the proportion of their alcohol use from beer was higher than among men in earlier cohorts. Over the same time period, the proportion of alcohol from wine showed a gradual increase with age for all birth cohorts. There was <15% of alcohol from wine for men before age 40 years, but it increased to more than 25% in later years. Between ages 40 and 60 years, men in the youngest birth cohort consumed a higher proportion of their alcohol from wine than did the earlier birth cohorts. There was less of an effect of age on the proportion of alcohol from liquor for all cohorts. However, the proportion of alcohol from liquor was lower at most ages for individuals born in the younger birth cohorts.

For women, the percentage of alcohol derived from beer was much lower than it was for men, but a similar decrease with age was seen for all birth cohorts. Women consumed a higher percentage of alcohol from wine and, as in men, the percentage increased with age for all birth cohorts. The highest percentage from wine was seen for the youngest birth cohort. For liquor, the pattern with age was similar for women as for men, without substantial effects for the older birth cohorts. Women in the youngest birth cohort, however, consumed less of their alcohol from liquor, and showed a steady decrease with age.

For both men and women, there was a steady increase in abstinence with age for all birth cohorts. By age 80 years, more than 40% of men and 60% of women reported no alcohol consumption. For the earliest birth cohort (born before 1900), women in their 50s and 60s were more likely to be abstainers than were women in later cohorts. In the 3 younger birth cohorts, the prevalence ratios (PR) of abstinence among men were 0.75 and 0.73, respectively, for participants born from 1900-1919 and 1920-1939 in comparison with those born in 1940-1959. In women, in the later cohorts, the corresponding figures were 0.80 and 1.01, respectively.

For the birth cohorts with available data, among individuals reporting any alcohol intake, the proportion reporting "moderate" amounts of alcohol increased with age in men (Figure 2A), but somewhat less so in women (Figure 2B). Further, where the ages overlapped, for both men and women, the younger birth cohort tended to show a higher prevalence of such a drinking pattern than the older one. For example, for men between ages 40 and 55 years, the age-adjusted PR of "moderate" drinking among alcohol consumers in the youngest cohort was 1.21 (95% confidence interval [CI], 1.12-1.32) in comparison with subjects born from 1920-1939; for women in these cohorts, the PR was 1.18 (95% CI, 1.10-1.27). Between ages 60 and 75 years, for men and women combined, the age-adjusted PR of "moderate" alcohol drinking was 1.17 (95% CI, 1.07-1.28) when comparing individuals born from 1920-1939 with those born from 1900-1919 ($P < .013$).

In contrast, the prevalence of "heavy" drinking for the 3 more recent birth cohorts was lower in the younger cohorts

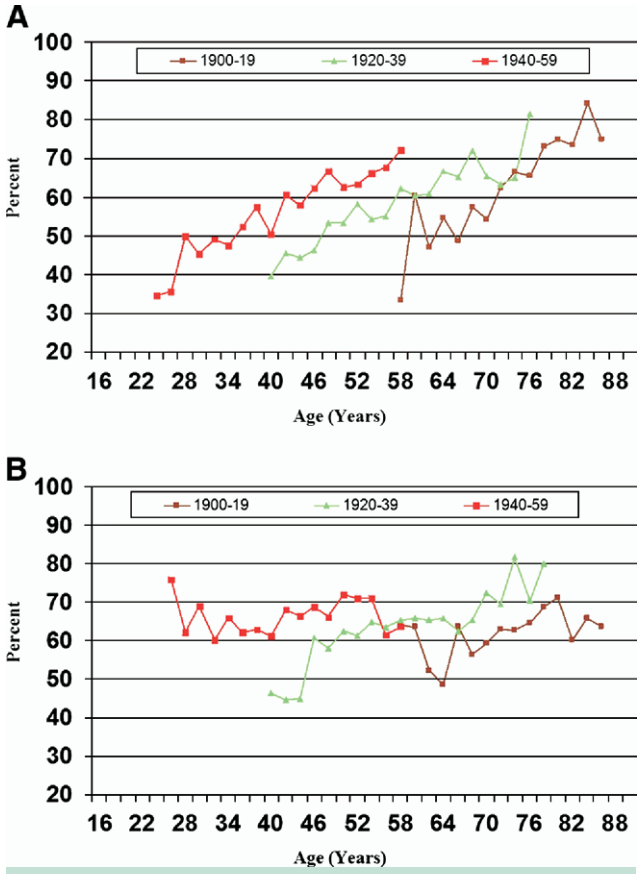


Figure 2 **(A)** Proportion of male participants reporting any alcohol intake whose intake was considered "moderate" (mean ≤24 g of alcohol per day and no heavy drinking episodes), by age and birth cohort. **(B)** Proportion of female participants reporting any alcohol intake whose intake was considered "moderate" (mean ≤12 g of alcohol per day and no heavy drinking episodes), by age and birth cohort.

than in the older cohorts for both men (Figure 3A) and women (Figure 3B). For men, the proportion with "heavy" drinking tended to decrease with age, although the pattern was less clear among women.

As shown in Table 2, the cumulative incidence of evidence of an alcohol use disorder between ages 40 and 79 years showed a slight increase from the cohort born between 1900 and 1919 to the later birth cohort; the overall cumulative incidence for the combined cohorts was 12.8% for men and 3.8% for women. When the analysis included individuals with *either* a code for an alcohol use disorder *or* a keyword indicating a possible disorder, the cumulative incidence increased very little: 13.8% for men and 4.4% for women.

DISCUSSION

Using repeatedly collected data on alcohol use and consequences over 50 years in the Framingham Heart Study, we found that the average amount of alcohol consumed by drinkers decreased across birth cohorts. The proportion of individuals reporting "moderate" consumption increased and the proportion reporting "heavy" intake decreased. Further, the percentage of the alcohol intake from beer decreased and the percentage from wine increased across birth cohorts.

The effects of age were similar for most cohorts. The average amount of alcohol was highest between ages 30 and 50 years, then decreased with age. The percentage of alcohol derived from beer decreased and that derived from wine increased with age; the proportion from liquor showed little change with age. Proportion of abstinence increased markedly with age. Among participants in the older cohorts where data were available, the cumulative incidence of identifying an alcohol use disorder between ages 40 and 79 years was about 3 times higher for men than for women, but showed little change across cohorts.

Several characteristics of this study are noteworthy. First, unlike cross-sectional surveys that have described patterns of alcohol consumption, in this study, data on alcohol were collected repeatedly in the same subjects fol-

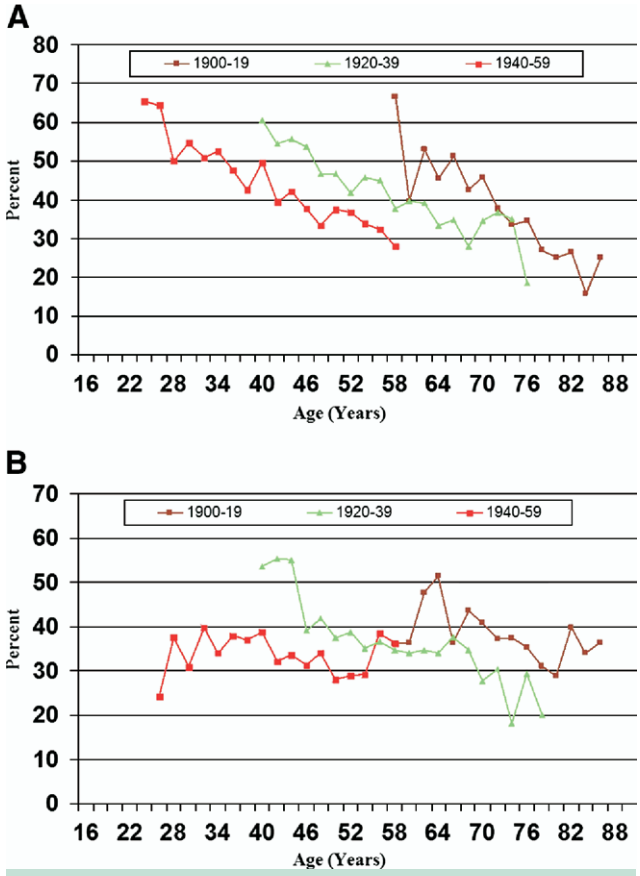


Figure 3 **(A)** Proportion of male participants reporting any alcohol intake whose intake was considered "heavy" (mean >24 g of alcohol per day or heavy drinking episodes), by age and birth cohort. **(B)** Proportion of female participants reporting any alcohol intake whose intake was considered "heavy" (mean >12 g of alcohol per day or heavy drinking episodes), by age and birth cohort.

Table 2 Cumulative Incidence of an Alcohol Use Disorder (ICD-9 Code for 100% Attributable Disease) Between Ages 40 and 79 Years, by Sex and Birth Cohort from The Framingham Heart Study, Framingham, MA, 1948-1975

	Men		Women	
Birth Cohort	Number	Risk and 95% CI (%)	Number	Risk and 95% CI (%)
1900-1919	694	11.7 (9.1-13.5)	768	3.5 (2.3-4.5)
1920-1939	902	13.6 (10.7-15.9)	904	4.0 (2.6-5.3)
Total	1596	12.8 (10.9-14.3)	1672	3.8 (2.9-4.6)

CI = confidence interval.

lowed for long periods of time. Secondly, as almost 90% of the participants in the Original Cohort and more than 16% of those in the Offspring Cohort had died by the end of the follow-up in these analyses, and the average age of survivors was 86.8 years for the cohort and 61.6 years for the offspring, the data tend to reflect the lifetime exposure to alcohol for a substantial proportion of the sample. Furthermore, because individuals in the current analysis were born over many decades (ranging from those born before 1900 to those born in 1940-1959), we have been able to describe secular changes in alcohol use over a large proportion of the 20th century, and to separate the effects of age from those of birth cohort.

Similar questions were asked to assess alcohol consumption at each examination, and we took into consideration secular changes in the alcohol content of specific alcoholic beverages as well as changes in the average serving sizes of drinks. Because reports of intake at repeated examinations of individuals were generally very stable, we believe that the overall estimates we report are unlikely due to measurement error. Further, the reported intake has been shown to relate to measured high-density lipoprotein levels in studies that included these subjects.[22]

Interestingly, cigarette smoking, another lifestyle factor, showed a similar trend of decrease across birth cohorts. Smoking rates decreased after the 1970s, and the percentage of current cigarette smokers was lower among the Offspring Cohort than among the Original Cohort.[23] Participants in the Framingham Study have been found to have values of lipids and other characteristics that reflect those of non-Hispanic Whites in the United States;[24] thus, we believe that the secular trends of alcohol use and misuse that we describe among the participants probably mirror national patterns.

Limitations

Our study has some limitations as well. First, the estimates of alcohol intake were all based on self report, and sought data on the subject's usual intake over time, which may be less accurate than estimates based on records kept of alcohol consumption. Because ethanol content and serving size changed during the study period, we developed a conversion formula to account for such changes; however, when we repeated our analyses excluding data collected at the earlier examinations, we found no appreciable differences. Secondly, an important measure of alcohol intake is the drinking pattern, and such information was not systematically recorded for some of the earlier examinations. When such data were collected, we used a commonly used definition for "heavy episodic drinking" based on the reported limit for number of drinks at one period of time. Also, our data are from mainly middle-aged and older white adults; we have limited data on younger individuals and none on minorities. Because many participants did not enter the Framingham Study until they were in their 50s, we are unable to calculate the lifetime risk of developing an alcohol use disorder. Our estimates of risk of developing such a disorder are limited to those not showing such alcohol problems in medical records before age 40 years. However, because most alcohol use disorders develop before age 40 years, it is likely that the disorders first recognized in these records as incident represent disorders that had their initial symptoms in young adulthood.

The Framingham Study did not have specific questions to solicit the evidence of an alcohol use disorder; thus, our estimate on the occurrence of such a disorder was based on the appearance of a relevant diagnosis in the medical records. It is quite possible that we may have underestimated the prevalence of these disorders, even after adding those with possible disorders from key words within the text of records, because many such consequences do not come to medical attention.

Grant et al[15] showed that "alcohol abuse" increased between cross-sectional surveys in 1991-1992 and 2001-2002, whereas "alcohol dependence" showed a decrease. Of note, the prevalence of current alcohol use disorder in their sample (12% for men, 5% for women) was similar to the proportions identified in our study. However, it is difficult to compare the cumulative rates of alcohol use disorders among the same subjects over the very long period of time that we report with the prevalence from cross-sectional surveys among different individuals.[15,16]

CONCLUSION

These analyses provide data that may be useful for groups making recommendations about drinking and setting alcohol policy. Research continues to support potentially beneficial effects of moderate drinking on cardiovascular disease

and other diseases of aging, as well as adverse health and social effects of heavy alcohol use. The findings in this study may be considered encouraging in many ways: the average amount of alcohol has decreased in more recently born cohorts, the percentage of the population exhibiting "moderate" alcohol intake has been increasing steadily, and the percentage reporting "heavy" drinking has decreased over time. Further, with greatly increasing numbers of elderly people, the trends for less "heavy" drinking with advancing age for all birth cohorts also is encouraging. While these data suggest the development of more favorable patterns of alcohol consumption over the latter part of the 20th century, they also show that, at the same time, the cumulative incidence of alcohol use disorders has not shown a decrease, and continuing efforts at preventing them are warranted.

References

1. Bloss G. Measuring the health consequences of alcohol consumption: current needs and methodological challenges. *Dig Dis*. 2005;23:162-169.
2. Dawson DA, Grant BF, Li TK. Quantifying the risks associated with exceeding recommended drinking limits. *Alcohol Clin Exp Res*. 2005; 29:902-908.
3. Saitz R. Clinical practice. Unhealthy alcohol use. *N Engl J Med*. 2005;352:596-607.
4. Ellison RC, Zhang Y, McLennan CE, Rothman KJ. Exploring the relation of alcohol consumption to risk of breast cancer. *Am J Epidemiol*. 2001;154:740-747.
5. McPherson K. Moderate alcohol consumption and cancer. *Ann Epidemiol*. 2007;17(5 Suppl):S46-S48.
6. Rimm EB, Giovannucci EL, Willett WC, et al. Prospective study of alcohol consumption and risk of coronary disease in men [see comments]. *Lancet*. 1991;338:464-468.
7. Walsh CR, Larson MG, Evans JC, et al. Alcohol consumption and risk for congestive heart failure in the Framingham Heart Study. *Ann Intern Med*. 2002;136:181-191.
8. Mukamal K. Alcohol intake and noncoronary cardiovascular diseases. *Ann Epidemiol*. 2007;17(5 Suppl):S8-S12.
9. Koppes LL, Dekker JM, Hendriks HF, et al. Moderate alcohol consumption lowers the risk of type 2 diabetes: a meta-analysis of prospective observational studies. *Diabetes Care*. 2005;28:719-725.
10. Stampfer MJ, Kang JH, Chen J, et al. Effects of moderate alcohol consumption on cognitive function in women. *N Engl J Med*. 2005; 352:245-253.
11. Mukamal KJ, Robbins JA, Cauley JA, Kern LM, Siscovick DS. Alcohol consumption, bone density, and hip fracture among older adults: the cardiovascular health study. *Osteoporos Int*. 2007;18:593-602.
12. Lang I, Guralnik J, Wallace RB, Melzer D. What level of alcohol consumption is hazardous for older people? Functioning and mortality in U.S. and English national cohorts. *J Am Geriatr Soc*. 2007;55:49-57.
13. Breslow RA, Smothers BA. Drinking patterns and body mass index in never smokers: National Health Interview Survey, 1997-2001. *Am J Epidemiol*. 2005;161:368-376.
14. Rimm EB, Moats C. Alcohol and coronary heart disease: drinking patterns and mediators of effect. *Ann Epidemiol*. 2007;17(5 Suppl): S3-S7.
15. Grant BF, Dawson DA, Stinson FS, et al. The 12-month prevalence and trends in DSM-IV alcohol abuse and dependence: United States, 1991-1992 and 2001-2002. *Drug Alcohol Depend*. 2004;74:223-234.
16. Breslow RA, Guenther PM, Smothers BA. Alcohol drinking patterns and diet quality: the 1999-2000 National Health and Nutrition Examination Survey. *Am J Epidemiol*. 2006;163:359-366.
17. Newes-Adeyi G, Chen CM, Williams GD, Faden VB. *Surveillance Report #74: Trends in Underage Drinking in the United States, 1991-2003*. Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism; 2005.
18. Yoon YH, Yi HY. *Surveillance Report #75: Liver Cirrhosis Mortality in the United States, 1970-2003*. Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism; 2006.
19. Yi HY, Chen CM, Williams GD. *Surveillance Report #76: Trends in Alcohol-Related Fatal Traffic Crashes, United States, 1977-2004*. Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism; 2006.
20. Castelli WP. The Framingham Heart Study—thirty years. *Med Sect Proc*. 1987;103-114.
21. Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*. 1986;42:121-130.
22. Djousse L, Levy D, Herbert AG, et al. Influence of alcohol dehydrogenase 1C polymorphism on the alcohol-cardiovascular disease association (from the Framingham Offspring Study). *Am J Cardiol*. 2005; 96:227-232.
23. Wilson PWF, D'Agostino RB, Kannel WB. Secular trends in cardiovascular risk factor levels: the Framingham perspective. In: Jacotot B, Mathé D, Fruchart J-C, eds. *Atherosclerosis XI*. Singapore: Elsevier; 1998:659-663.
24. Vasan RS, Sullivan LM, Wilson PW, et al. Relative importance of borderline and elevated levels of coronary heart disease risk factors. *Ann Intern Med*. 2005;142:393-402.

BMC Health Services Research

Research article

Healthcare costs and utilization for Medicare beneficiaries with Alzheimer's

Yang Zhao*[1], Tzu-Chun Kuo[2], Sharada Weir[3,6], Marilyn S Kramer[4,6] and Arlene S Ash[2,5]

Address: [1]Eli Lilly and Company, Indianapolis, IN, USA, [2]DxCG, Inc, Boston, MA, USA, [3]Center for Health Policy and Research/Department of Family Medicine and Community Health, University of Massachusetts Medical School, Shrewsbury, MA, USA, [4]Partnership for Healthcare Excellence, Boston, MA, USA, [5]Health Care Research Unit, Section of General Internal Medicine, Department of Medicine, Boston University School of Medicine, Boston, MA, USA and [6]Formerly of DxCG, Inc, Boston, MA, USA

Email: Yang Zhao* - zhao_yang_yz@lilly.com; Tzu-Chun Kuo - tzu.kuo@dxcg.com; Sharada Weir - sharada.weir@umassmed.edu; Marilyn S Kramer - mkramer@healthexcellence.net; Arlene S Ash - aash@bu.edu

* Corresponding author

Published: 22 May 2008

BMC Health Services Research 2008, **8**:108 doi:10.1186/1472-6963-8-108

Received: 9 November 2007
Accepted: 22 May 2008

This article is available from: http://www.biomedcentral.com/1472-6963/8/108

Abstract

Background: Alzheimer's disease (AD) is a neurodegenerative disorder incurring significant social and economic costs. This study uses a US administrative claims database to evaluate the effect of AD on direct healthcare costs and utilization, and to identify the most common reasons for AD patients' emergency room (ER) visits and inpatient admissions.

Methods: Demographically matched cohorts age 65 and over with comprehensive medical and pharmacy claims from the 2003–2004 MEDSTAT MarketScan® Medicare Supplemental and Coordination of Benefits (COB) Database were examined: 1) 25,109 individuals with an AD diagnosis or a filled prescription for an exclusively AD treatment; and 2) 75,327 matched controls. Illness burden for each person was measured using Diagnostic Cost Groups (DCGs), a comprehensive morbidity assessment system. Cost distributions and reasons for ER visits and inpatient admissions in 2004 were compared for both cohorts. Regression was used to quantify the marginal contribution of AD to health care costs and utilization, and the most common reasons for ER and inpatient admissions, using DCGs to control for overall illness burden.

Results: Compared with controls, the AD cohort had more co-morbid medical conditions, higher overall illness burden, and higher but less variable costs ($13,936 s. $10,369; Coefficient of variation = 181 vs. 324). Significant excess utilization was attributed to AD for inpatient services, pharmacy, ER visits, and home health care (all $p < 0.05$). In particular, AD patients were far more likely to be hospitalized for infections, pneumonia and falls (hip fracture, syncope, collapse).

Conclusion: Patients with AD have significantly more co-morbid medical conditions and higher healthcare costs and utilization than demographically-matched Medicare beneficiaries. Even after adjusting for differences in co-morbidity, AD patients incur excess ER visits and inpatient admissions.

Background

Alzheimer's disease (AD) is a progressive, irreversible neurodegenerative disorder with high social and economic costs. Currently, an estimated 5.1 million Americans have AD, 4.9 million of them over the age of 65 [1]. Alzheimer's disease affects 13% of people over age 65 and nearly half of those over age 85, accounting for 50 to 70% of all dementia cases [1]. By 2050, 11.6 to 16 million Americans may have AD [2]. With the expected increase in AD cases, medical costs for Medicare beneficiaries with AD are expected to increase from $91 billion in 2005 to $160 billion in 2010 [3]. Understanding what contributes to health care costs and utilization among AD patients should help health plans develop effective disease management protocols.

Prior studies [4-16] on costs and utilization associated with AD in the US have several limitations: 1) most use data collected prior to 2000 which do not capture current treatment patterns [4-12,15,16]; 2) several "contaminate" their definitions of AD cohorts by including diagnosis codes of non-AD dementias in their claims data analyses [4-6,8,12]; 3) several present only aggregated total cost data or omit pharmacy costs [4,9,10,12]; and 4) most rely on the presence of, at most, a small subset of co-morbidities to account for the effect of differences in disease burden between cases and controls [4-16].

In addition, no previous research has explored which medical conditions lead to inpatient admissions and emergency room (ER) use for patients with AD or estimated how AD affects such utilization beyond what would be expected based on the presence of co-morbidities.

In this study we sought to identify differences in direct healthcare costs and utilization, and common reasons for ER visits and inpatient admissions between Medicare beneficiaries with an AD diagnosis and controls, after comprehensively adjusting for the presence of other co-morbidities. Using a large, US administrative claims database, we examined 2004 direct healthcare costs and utilization for individuals aged 65 and above in each cohort with comprehensive medical (including claims paid by Medicare) and pharmacy claims. Illness burden was measured by a comprehensive disease classification and scoring system and used to produce estimates of the marginal effect of AD on inpatient, ER, pharmacy and other utilization and costs. Reasons for ER visits and inpatient admissions were also examined for both AD and control cohorts, and logistic regression was used to assess the contribution of AD to the most common reasons for ER visits and hospitalizations controlling for differences in overall illness burden. Findings from this research may help healthcare providers and health plans in the US develop protocols to better manage patients with AD.

Table 1: Demographics, Health Plan Enrollment, and Illness Burden

	AD Cohort	Control Group
	N = 25,109	N = 75,327
Characteristic		
Age, Mean (SD)*	80.1 (6.5)	80.1 (6.6)
Age, n (%)		
65 – 69	1,529 (6.1)	4,613 (6.1)
70 – 74	3,482 (13.9)	10,330 (13.7)
75 – 79	6,345 (25.6)	19,641 (26.1)
80 – 84	7,326 (29.2)	21,471 (28.5)
85 – 89	4,654 (18.5)	13,428 (17.8)
90+	1,683 (6.7)	5,844 (7.8)
Female, N (%)	15,473 (61.6)	47,082 (62.5)
Regional distribution, N (%)		
Northeast	2,986 (11.9)	10,783 (14.3)
North Central	8,710 (34.7)	25,725 (34.2)
South	9,178 (36.6)	24,028 (31.9)
West	4,185 (16.7)	14,534 (19.3)
Unknown	50 (0.2)	257 (0.3)
Months of health plan enrollment in 2004, Mean (SD)	11.1 (2.51)	11.2 (2.34)
Enrolled all 12 months of 2004, N (%)	21,380 (85.2)	65,949 (87.6)
Number of Comorbidities, Mean (SD)†§	8.1 (5.5)	6.5 (5.0)
Illness burden score, Mean (SD)‡§	1.23 (0.80)	1.04 (0.77)

* Age as of 01 Jan 2003 displayed as mean ± standard deviation (SD)
† Diagnostic Cost Group (DCG) Condition Categories (CCs), excluding AD.
‡ Prospective Relative Risk Score (RRS) excluding AD, normalized to 1.0 for the eligible study population.
§ Statistically significant differences between the two groups at $p < 0.001$.

Methods

Study Sample

Data were obtained from the MEDSTAT MarketScan® Medicare Supplemental and Coordination of Benefits (COB) Database for 2003 and 2004. These data contain information about Medicare beneficiaries across the US, age 65 and older, with employer-sponsored supplemental insurance including pharmacy benefits. The eligible study population was comprised of 627,775 individuals with 12 months of enrollment in 2003, at least one month of enrollment in 2004, and comprehensive inpatient, outpatient, and pharmacy claims. Coverage for skilled nursing facility care is limited under Medicare, and generally long term nursing home care is not covered by employer-sponsored supplemental insurance. Thus, the cost and utilization data presented here do not include care provided in skilled nursing facilities and nursing homes. Because data used in this study were purchased from a third party which had removed identifying information prior to its release, institutional review board (IRB) and similar approvals were neither needed nor sought.

From this population we selected 25,109 individuals with AD and three times as many demographically-matched controls. The AD cohort contained all patients with at least one non-laboratory claim with an AD diagnosis (ICD-9-CM code 331.0) or at least one filled prescription for an AD-specific medication (tacrine, donepezil, galantamine, rivastigmine, or memintine) in 2003. The control cohort excluded both those with AD using the above criteria and those with any other form of dementia (ICD-9-CM codes 290.xx, 294.1x, 294.8x, 331.1x-331.9x, 797.xx) recorded at any time in either 2003 or 2004. Because a diagnosis of AD is not always accurate, it is likely that the AD cohort includes some number of people with dementia arising from other causes, such as vascular dementia.

We used propensity scores to construct the demographically matched control cohort. First we predicted the probability of "propensity to have" AD using logistic regression. Explanatory variables included: 1) age as of January 1, 2003, 2) age squared, 3) gender, 4) geographic region, and 5) months of health plan enrollment in 2004. The data were ranked by the probability of having AD (from highest to lowest), and then partitioned into twenty quantiles (with 5% of the population in each). Within each 5% quantile, eligible individuals were randomly selected to match each individual in the AD cohort using a 3 to 1 ratio.

Controlling for Illness Burden

We measured overall illness burden using DxCG, Inc.'s RiskSmart™ software, which includes the Diagnostic Cost Groups (DCGs) model. This validated, diagnosis-based classification system [17] organizes over 15,000 ICD-9-CM codes into 781 highly homogeneous clinical categories (DxGroups), which are further clustered into 184 Condition Categories (CCs) that encompass similar medical conditions with similar expected costs [18]. The software further organizes the CC information for each person into a 184-variable vector of "hierarchical CCs" (HCCs), where the presence of a more serious manifestation of a disease causes clinically-less-relevant conditions to be "zeroed out." For example, "chronic obstructive lung disease" dominates "cold" [18].

Table 2: 2004 Annualized Health Care Costs and Utilization by Type of Service

	AD			Controls			Contribution of AD†	
	% of Cohort Using	Mean Visit Days	Mean Cost $ (CV)*	% of Cohort Using	Mean Visit Days	Mean Cost $ (CV)*	Visit Days†	Cost $ †
Inpatient	30	3.38	5,094 (419)	20	1.93	4,014 (753)	1.14‡	671‡
Pharmacy	94	--	4,056 (77)	85	--	2,169 (174)	--	1,711‡
Hospital outpatient	61	2.61	1,252 (298)	58	2.51	1,412 (384)	-0.30‡	-366‡
Physician office	79	8.46	1,100 (237)	81	9.97	1,508 (270)	-2.86‡	-648‡
Emergency room	41	1.04	335 (280)	27	0.64	196 (530)	0.28‡	107‡
Home health	7	0.24	32 (1,001)	4	0.15	27 (3,191)	0.05‡	0.37
Other	56	--	2,068 (353)	37	--	1,043 (571)	--	789‡
Total Utilization	97	--	13,936 (181)	91	--	10,369 (324)	--	2,307‡

* CV = Coefficient of Variation = 100*SD/Mean

† Coefficient of the AD indicator from the weighted least square regression for annualized costs, or utilization within each setting controlling for overall illness burden (RRS excluding AD and RRS squared).

‡ Statistically significant effects attributed to AD at $p < 0.05$.

We predicted 2004 health care costs for the eligible study population using HCCs from the 2003 diagnoses, age, and sex. The CC for AD was excluded from the estimation so that the AD and control cohorts would have comparable predicted costs after controlling for demographics and all co-morbidities other than AD in 2003.

Predicted costs were then normalized (multiplied by the appropriate constant) to create prospective relative risk scores (RRSs) that average 1.00 in the eligible study population.

Analyses

Demographics such as age, gender, and region were used to characterize the AD cohort and controls. We also calculated mean duration of health plan enrollment in 2004, the number of unique co-morbid conditions, and the RRS for non-AD illness burden.

Health care costs and utilization were calculated for 2004 by place of service, including physician office visits, outpatient hospital services, ER visits, inpatient services, home health care, and pharmacy. All costs included deductibles, copayments, coinsurance, and coordination-of-benefits payments. For partial year enrollees, utilization and expenditures were annualized (actual totals divided by percentage of year enrolled). Total health care costs were compared using the mean, computed t-tests of the differences in means, standard deviation (SD), cost dispersion as measured by the coefficient of variation (CV) and cost distribution by place of service. Health care utilization was evaluated by examining percentages of users, the mean cost per member per year and, where applicable, the number of encounters (visits) per member per year.

Regression was used to estimate AD's independent effect on overall costs and utilization. Specifically, Weighted Least Squares (WLS) regressions were used to assess the marginal contribution of AD to 2004 costs and health care utilization, weighted by the fraction of time enrolled in 2004. Explanatory variables included: 1) an indicator variable (0/1) identifying individuals in the AD cohort; 2) the RRS, as a measure of total non-AD illness burden; and 3) RRS squared. The coefficient associated with the AD indicator measures the extent to which AD contributes to excess healthcare costs or utilization after controlling for differences in overall illness burden.

The most common reasons for inpatient admissions and for ER visits were identified by mapping the first diagnosis from each claim into a DxGroup category. We also compared the prevalence per 10,000 persons of ER visits and inpatient admissions between the AD cohort and controls. We assessed the marginal contribution of AD to ER visits and inpatient admissions controlling for difference in overall illness burden via logistic regressions using the same predictors in WLS regressions as noted above.

All analyses were performed using SAS software (version 9.1, SAS, Cary, NC).

Table 3: Top 10 Reasons for ER Visits by Cohort

	AD	Control	Odds Ratio*
	Rate per 10,000 (Rank within cohort)†		
Contusion/superficial injury	**679 (1)**	**270 (2)**	2.23‡
Chest pain	**480 (2)**	**353 (1)**	1.16‡
Syncope and collapse	**333 (3)**	**115 (10)**	2.64‡
Open wound except eye and lower arm	**316 (4)**	**115 (11)**	2.54‡
Cystitis, other urinary tract infections	**306 (5)**	103 (13)	2.69‡
Other general symptoms	**300 (6)**	**130 (7)**	2.07‡
Other and unspecified pneumonia	**288 (7)**	**145 (5)**	1.74‡
Abdominal/pelvis symptoms	**279 (8)**	**219 (3)**	1.10
Stupor/altered consciousness/trans global amnesia/febrile convulsions	**264 (9)**	41 (37)	5.85‡
Disorders of fluid/electrolyte/acid-base balance, e.g., dehydration	**252 (10)**	94 (16)	2.36‡
Heart failure	192 (14)	**181 (4)**	0.87‡
Stomach/intestinal disorders/symptoms, except obstruction, ulcer, and hemorrhage	223 (12)	**136 (6)**	1.43‡
Arthropathy/joint disorders, derangements, joint pain/stiffness, excluding gout	193 (13)	**122 (8)**	1.49‡
Nonspecific backache and other back/neck pain/disorders	141 (20)	**121 (9)**	1.05
Any ER visit	10,413	5,733	1.74‡

* Odds ratio of the AD indicator from the logistic regression predicting any ER visit controlling for overall illness burden (RRS excluding AD and RRS squared).

† 10,000 times the number of ER visits divided by number of individuals in the cohort. Bold text designates reason in the top ten.

‡ Statistically significant AD effect at $p < 0.05$.

Results

As constructed, the control group did not differ from the AD cohort with respect to age, sex, regional distribution, or mean length of eligibility in 2004 (Table 1). However, the AD cohort had more co-morbidities (mean of 8.1 CCs vs. 6.5, $p < 0.001$), and a greater burden of non-AD illness (mean RRS of 1.23 vs. 1.04, $p < 0.001$).

Annualized costs and utilization by place of service are summarized in Table 2. The vast majority of AD patients (97%) and controls (91%) used some healthcare services in 2004. Rates of ER visits (41% vs. 27%), inpatient hospital stays (30% vs. 20%), and home health care (7% vs. 4%) were about 50% higher for AD patients than for controls. Controlling for overall illness burden, the excess utilization attributed to AD for inpatient services, ER visits, and home health care were all significant ($p < 0.05$). However, the AD cohort used fewer physician office visits and outpatient hospital services (both $p < 0.05$). Differences in costs attributed to AD were statistically significant ($p < 0.05$) for all categories except home health care. Spending in the AD cohort was higher for all but two categories of services (outpatient services and office visits) compared to controls, but AD spending was always less variable (coefficient of variation (CV) was lower). Excess pharmacy costs associated with AD were $1,711, more than twice that of any other expense category ($p < 0.05$).

The 10 most common reasons for ER visits among individuals belonging to the AD or control cohorts, and their rates and comparative prevalence, are shown in Table 3. The AD and control cohorts shared 6 of their 10 most common reasons, and their top two reasons (contusion/superficial injury and chest pain) were the same. The AD cohort had higher raw ER use rates than controls for all 14 reasons listed, and significantly higher risk-adjusted use rates for 11 of them with odds ratios (ORs) ranging as high as 5.85 for stupors and other states of altered consciousness. Only one of the 14 reasons shown (heart failure) had risk-adjusted ER use rate lower for AD patients than for controls (OR = 0.87, $p < 0.05$). The odds of an individual in the AD cohort having an ER visit for any reason was 74% greater than for controls.

Analogous to the previous table, Table 4 compares utilization for the top 10 most common reasons for inpatient admission for each cohort. Only 3 of the top reasons (hip fracture, other and unspecified pneumonia, and cystitis/other urinary tract infections) were shared. Even after controlling for illness burden, patients in the AD cohort were more likely to be hospitalized for most of the reasons listed, including pneumonia, infections, syncope and hip fracture. However, AD patients were less likely to be admitted to the hospital due to heart failure or coronary atherosclerosis and other coronary ischemic heart disease (all $p < 0.05$) compared with those in the control cohort.

Table 4: Top 10 Reasons for Inpatient Admission by Cohort

	AD	Control	Odds Ratio*
	Rate per 10,000 (Rank within cohort)†		
Other and unspecified pneumonia	**229 (1)**	**133 (2)**	1.50‡
Femoral (hip) fracture	**209 (2)**	**88 (4)**	2.32‡
Cystitis, other urinary tract infections	**161 (3)**	42 (13)	3.46‡
Heart failure	**150 (4)**	**158 (1)**	0.78‡
Cerebral degeneration/Alzheimer's disease	**142 (5)**	-	infinite
Disorders of fluid/electrolyte/acid-base balance, e.g., dehydration	**118 (6)**	49 (11)	2.16‡
Septicemia (blood poisoning)/shock	**118 (7)**	14 (39)	2.77‡
Syncope and collapse	**100 (8)**	33 (30)	2.85‡
Aspiration pneumonia	**100 (9)**	18 (16)	5.36‡
Pre-cerebral or cerebral arterial occlusion with infarction	**84 (10)**	48 (12)	1.74‡
Coronary atherosclerosis and other chronic ischemic heart disease (CAD)	70 (13)	**111 (3)**	0.53‡
Acute myocardial infarction, initial episode of care	83 (11)	**81 (5)**	0.92
Atrial arrhythmia	56 (16)	**62 (6)**	0.83
Osteoarthritis of lower leg (knee)	27 (29)	**62 (7)**	0.39‡
Emphysema/chronic bronchitis, 18+	65 (14)	**55 (8)**	1.02
Gastrointestinal hemorrhage, except peptic ulcer and anal/rectal	78 (12)	**55 (9)**	1.22‡
Chest pain	54 (17)	**50 (10)**	0.94
Any inpatient admission	3,796	2,408	1.55‡

* Odds ratio of the AD indicator from the logistic regression predicting any inpatient admission controlling for overall illness burden (RRS excluding AD and RRS squared).

†10,000 times the number of inpatient admissions divided by number of individuals in the cohort. Bold text designates reason in the top ten.

‡ Statistically significant AD effect at $p < 0.05$.

The only other reason for hospital admission that was less likely in the AD cohort was for osteoarthritis of lower leg (knee), typically related to knee surgery. The odds for an inpatient admission were 55% greater for AD patients than for controls.

Discussion

Our study is the first to examine the most common medical conditions among AD patients that lead to inpatient admissions and ER use, and to contrast AD patients' co-morbid diseases and utilization with those of a demographically-matched control group. Earlier studies accounted for just a few co-morbidities in seeking to isolate AD's influence on costs and utilization. Using a comprehensive co-morbidity assessment for AD patients and a demographically-matched control group, we have achieved credible estimates of the independent effect of AD on healthcare costs and utilization. Individuals in the AD cohort had more unique co-morbid medical conditions and higher overall illness burden than those in the control cohort. Use of services was greater for AD patients, with far more inpatient, ER, and home health encounters. Mean excess cost attributed to AD, even after controlling for the greater overall burden of illness, was $2,307.

The AD cohort had considerably higher pharmacy costs [7,8,13,14] and total health care costs [4,7,8,11-13] than seen in previous studies. The findings establish the need to better understand pharmacy management practices for AD patients given pharmacy's large contribution to their elevated costs. Compared with demographically-matched controls, AD patients had significantly higher but less variable costs. So although patients with AD were costly, their costs were more predictable than those in the control cohort.

In our study, AD patients on average had longer hospital days and more visits than controls for all utilization categories except physician office visits. Overall prevalence rates for ER visits and inpatient admissions were significantly higher for AD patients. Closely managing hospitalization as well as ER visits may have significant impact on health care resource use in AD.

Alzheimer's disease complicates the management of an elderly population with significant co-morbid disease. Patient non-cooperation, inability to communicate, less frequent office visits, and caregiver burden may all contribute to "simple" medical problems escalating into hospital admissions or ER visits for reasons such as pneumonia, dehydration and septicemia. Dehydration, for example, may precipitate other medical problems, including cystitis, electrolyte imbalances, contusions, and hip fractures. However, it is unclear if dehydration requiring ER care or hospitalization is really more common in AD patients; an alternative explanation is that "dehydration" is used to code admissions requested by stressed caregivers in the absence of clear clinical symptoms.

Alzheimer's patients have many co-morbid medical problems and use multiple medications [14,19,20]; they may be prone to harm themselves [13]. The use of multiple medications raises the risk of adverse drug reactions and drug-drug interactions, and complicates medication compliance [21]. Polypharmacy, especially in the elderly population, is associated with adverse drug reactions [22-24], which occur in 5–10% of hospital inpatient admissions and increase hospital stays and costs [25]. All these factors, especially when combined with impaired cognition, could contribute to the observed increase in hospitalizations for hip fracture and syncope for AD patients. These findings suggest opportunities for improvement through case management to address AD patients' co-morbidities, specifically through medication reviews.

Differences in disease prevalence also lead to higher rates of hospitalization (most prominently, hospitalizations for AD itself). However, differences in co-morbidities do not explain AD patients' lower use of hospitalizations for heart problems since heart problems were similarly common in the two cohorts. This may be due to reduced awareness of (non-obvious) heart problems or because heart problems are treated less aggressively in AD patients. For example, ER visits for "chest pain" were more common in patients with AD compared to controls, although hospital admissions were less common.

Our study has several limitations. First, we examined Medicare-eligible individuals with employer-sponsored supplemental insurance, mostly from large companies whose active or former employees do not necessarily represent the general population of Medicare beneficiaries. This may contribute to the relatively low (4%) AD prevalence in this elderly cohort. Second, AD patients in our study were identified via diagnoses in administrative claims. Thus, some non-AD patients may be in our AD cohort (due to a false AD diagnosis) while some people with AD will be excluded (due to either a lack of any AD diagnosis or to misclassification, for example, as vascular dementia). Thirdly, the costs provided in our data do not capture care provided in skilled nursing facilities or nursing homes, therefore our analysis may underestimate the total direct healthcare costs of AD. Fourthly, our claims data do not have information on the duration or severity of AD, which is significantly related to healthcare cost and utilization [7,26]. Although we controlled for differences in overall co-morbidities in our analysis, we could not control for disease severity. Fifthly, our data do not include information on living situation (e.g., home versus institution), which may also affect healthcare costs and

utilization. Finally, although research has indicated that the indirect costs of AD are substantial [1], this study focused only on direct healthcare costs.

Conclusion

Compared to a demographically-matched control cohort, AD patients had significantly more co-morbid medical conditions and higher overall illness burden. Even after adjusting for differences in overall illness burden, people with AD incurred markedly more health care costs than their age-matched peers. The greater predictability of AD spending and more frequent, more costly, and different use of hospitals, suggests opportunities for improvement. Significantly increased financial exposure to AD with the expansion of prescription drug coverage under Medicare increases the pressure on health plans to develop more effective management protocols for AD patients.

Competing interests

YZ is employed by Eli Lilly and Company, and owns company stocks. T–CK and ASA are affiliated with DxCG, Inc. SW and MSK were previously affiliated with DxCG, Inc.

Authors' contributions

YZ conceived the study, participated in the design, and drafted the manuscript. T–CK performed the data analysis and helped draft the manuscript. MSK acquired the data, participated in the design and coordination of the study and helped draft the manuscript. SW participated in the study design, performed the data analysis and helped draft the manuscript. ASA participated in the design and coordination of the study and helped draft the manuscript. All authors read and approved the final manuscript.

Acknowledgements

This project was funded by Eli Lilly and Company. However, the sponsor had no role in the design, methods, data collection, analysis, or preparation of the manuscript. The interpretation and reporting of these data are the sole responsibility of the authors.

The authors would like to thank Tacey Ann Boucher, PhD, who is employed by Eli Lilly and Company, for her assistance in preparing and revising this manuscript.

References

1. **Alzheimer's disease facts and figures 2007. Statistical abstract of US data on Alzheimer's disease published by the Alzheimer's Association** [http://www.alz.org/national/documents/report_alzfactsfigures2007.pdf]
2. Hebert LE, Scherr PA, Bienias JL, Bennett DA, Evans DA: **Alzheimer disease in the U.S. population: prevalence estimates using the 2000 census.** *Archives of Neurology* 2003, **60(8):**1119-1122.
3. **Saving lives, saving money: dividends for Americans investing in Alzheimer's research. Report of the Lewin Group to the Alzheimer's Association. 2004** [http://www.alz.org/national/documents/report_savinglivessavingmoney.pdf]
4. Gutterman EM, Markowitz JS, Lewis B, Fillit H: **Cost of Alzheimer's disease and related dementia in managed Medicare.** *J Am Geriatr Soc* 1999, **47(9):**1065-1071.
5. Menzin J, Lang K, Friedman M, Neumann P, Cummings JL: **The economic cost of Alzheimer's disease and related dementias to the California Medicaid Program ("Medi-Cal") in 1995.** *Am J Geriatr Psychiatry* 1999, **7(4):**300-308.
6. Martin BC, Ricci JF, Kotzan JA, Lang K, Menzin J: **The net cost of Alzheimer disease and related dementia: a population-based study of Georgia Medicaid recipients.** *Alzheimer Dis Assoc Disord* 2000, **14(3):**151-159.
7. Fillit H, Hill JW, Futterman R: **Health care utilization and costs of Alzheimer's disease: the role of comorbid conditions, disease stage, and pharmacotherapy.** *Fam Med* 2002, **34(7):**528-535.
8. Hill JW, Futterman R, Duttagupta S, Mastey V, Lloyd JR, Fillit H: **Alzheimer's disease and related dementias increase costs of comorbidities in managed Medicare.** *Neurology* 2002, **58(1):**62-70.
9. Sloan FA, Taylor DH Jr: **Effect of Alzheimer's disease on the cost of treating other diseases.** *Alzheimer Dis Assoc Disord* 2002, **16(3):**137-143.
10. Weiner M, Powe NR, Weller WE, Shaffer TJ, Anderson GF: **Alzheimer's disease under managed care: implications from Medicare utilization and expenditure patterns.** *J Am Geriatr Soc* 1998, **46(6):**762-770.
11. Newcomer R, Clay T, Luxenberg JS, Miller RH: **Misclassification and selection bias when identifying Alzheimer's disease solely from Medicare claims records.** *J Am Geriatr Soc* 1999, **47(2):**215-219.
12. Richards KM, Shepherd MD, Crimson ML, Snyder EH, Jermain DM: **Medical services utilization and charge comparisons between elderly patients with and without Alzheimer's disease in a managed care organization.** *Clin Ther* 2000, **22(6):**775-791.
13. Frytak JR, Henk HJ, Zhao Y: **Health service utilization among Alzheimer's patients: evidence from managed care.** *Alzheimer's & Dementia* in press.
14. Joyce AT, Zhao Y, Bowman L, Flynn JA, Carter CT, Ollendorf TA: **Burden of illness among patients with Alzheimer's disease in a commercially-insured population.** *Alzheimer's & Dementia* 2007, **3(3):**204-210.
15. McCormick WC, Hardy J, Kukull WA, Bowen JD, Teri L, Zitzer S, Larson EB: **Healthcare utilization and costs in managed care patients with Alzheimer's Disease during the last few years of life.** *J Am Geriatr Soc* 2001, **49(9):**1156-1160.
16. Leibson C, Tomas O, O'Brien P, Waring S, Tangalos E, Hanson V, Plevack M, Kokmen E: **Use of physician and acuter care services by persons with and without Alzheimer's Disease: A population-based comparison.** *J Am Geriatr Soc* 1999, **47(7):**864-869.
17. Ash AS, Ellis RP, Pope GC, Ayanian JZ, Bates DW, Burstin H, Iezzoni LI, MacKay E, Yu W: **Using diagnoses to describe populations and predict costs.** *Health Care Financ Rev* 2000, **21(3):**7-28.
18. Zhao Y, Ash AS, Ellis RP, Ayanian JZ, Pope GC, Bowen B, Weyuker L: **Predicting pharmacy costs and other medical costs using diagnoses and drug claims.** *Medical Care* 2005, **43(1):**34-43.
19. Kuo T, Zhao Y, Weir S, Kramer MS, Ash AS: **Implications of comorbidity on costs for patients with Alzheimer's disease.** *Medical Care* in press.
20. Lacro JP, Jeste DV: **Physical comorbidity and polypharmacy in older psychiatric patients.** *Biological Psychiatry* 1994, **36(3):**146-152.
21. Stewart RB, Cooper JW: **Polypharmacy in the aged. Practical solutions.** *Drugs Aging* 1994, **4(6):**449-461.
22. Nolan L, O'Malley K: **Prescribing for the elderly: part 1. Sensitivity of the elderly to adverse drug reactions.** *J Am Geriatr Soc* 1988, **36(3):**142-149.
23. Colt HG, Shapiro AP: **Drug-induced illness as a cause for admission to a community hospital.** *J Am Geriatr Soc* 1989, **37(4):**323-326.
24. Pirmohamed M, James S, Meakin S, Green C, Scott AK, Walley TJ, Farrar K, Park BV, Breckenridge AM: **Adverse drug reactions as a cause for admission to hospital: prospective analysis of 18820 patients.** *Br Med J* 2004, **329(7456):**15-19.
25. Fattinger K, Roos M, Vergeres P, Holenstein C, Kind B, Masche U, Stocker DN, Braunschweig S, Kullak-Ublick GA, Galeazzi RL, Follath F, Gasser T, Meier PJ: **Epidemiology of drug exposure and adverse drug reactions in two Swiss departments of internal medicine.** *Br J Clin Pharmacol* 2000, **49(2):**158-167.

26. Feldman H, Gracon S: **Alzheimer's disease: symptomatic drugs under development.** In *Clinical Diagnosis and Management of Alzheimer's disease* Edited by: Serge Gauthier. London. Informa UK limited; 1996:236-259.

Pre-publication history

The pre-publication history for this paper can be accessed here:

http://www.biomedcentral.com/1472-6963/8/108/prepub