General Internal Medicine
Boston University School of Medicine
2004 Publications

1. **Ash AS**, **Carr PL**, Goldstein R, **Friedman RH**. Compensation and advancement of women in academic medicine: is there equity? *Ann Intern Med.* 2004;141:205-212.

2. Borzecki AM, Wong AT, Hickey EC, **Ash AS**, **Berlowitz DR**. Can we use automated data to assess quality of hypertension care? *Am J of Manag Care.* 2004;10:473-479.

3. **Carr PL**, Riccioti HA, **Freund KM**, Kahan S., eds. *In a page: OB/GYN & women's health*. Medford, Mass: Blackwell Publishing; 2004.

4. Charbonneau A, Rosen AK, Owen RR, Spiro A III, **Ash AS**, Miller DR, Kazis, L, Kader B, Cunningham F, **Berlowitz DR**. Monitoring depression care: in search of an accurate quality indicator. *Med Care.* 2004;42:522-531.

5. **Chuang CH**, Waldman LJ, **Freund KM**, **Ash AS**. Emergency contraception: prescribing practices of general internists compared with other primary care physicians. *Contraception.* 2004;69:43-45.

6. Conigliaro J, Madenwald T, Bryant K, Braithwaite S, Gordon A, Fultz SL, Maisto S, **Samet JH**, Kraemer K, Cook R, Day N, Roach D, Richey S, Justice A**.** The Veterans Aging Cohort Study: observational studies of alcohol use, abuse, and outcomes among human immunodeficiency virus-infected veterans. *Alcohol Clin Exp Res.* 2004;28:313-321.

7. **DeAlba I**, **Samet JH**, **Saitz R**. Burden of medical illness in drug- and alcohol-dependent persons without primary care. *Am J Addictions.* 2004;13:33-45.

8. DeMolles DA, Sparrow D, Gottlieb DJ, **Friedman RH**. A pilot trial of a telecommunications system in sleep apnea management. *Med Care.* 2004;42:764-769.

9. Ehrenstein V, Horton NJ, **Samet JH**. Inconsistent condom use among HIV-infected patients with alcohol problems. *Drug Alc Depend.* 2004;73:159-166.

10. **Farzanfar R**, Finkelstein J, **Friedman RH**. Testing the usability of two automated home-based patient-management systems. *J Med Syst.* 2004;28:143-153.

11. **Farzanfar R**, **Migneault J**, **Rubin A**, **Friedman R**, Marks L. Computer-based and live interviews on problem drinking: users' attitudes. *Medinfo*. 2004;11:956-959.

12. Flory J, Young-Xu Y, Gurol I, Levinsky N, **Ash A**, Emanuel E. Place of death: U.S. trends since 1980. *Health Aff.* 2004;23:194-200.

13. Frayne SM, **Freund KM**, Skinner KM, **Ash AS**, **Moskowitz MA**. Depression management in medical clinics: does healthcare sector make a difference? *Am J Med Qual.* 2004;19:28-36.

14. Frayne SM, Skinner KM, Lin H, **Ash AS**, **Freund KM**. Effect of patient gender on late life depression management. *J Women's Health.* 2004;13:919-925.

15. **Friedman RH**, Wahi-Gururaj S, Alpert J, Bauchner H, Culpepper L, Heeren T, Singer A. The views of U.S. medical school deans toward academic primary care. *Acad Med.* 2004; 79:1095-1102.

16. Krupitsky E, Zvartau E, Karandashova G, Horton NJ, Schoolwerth KR, Bryant K, **Samet JH**. The onset of HIV infection in the Leningrad region of Russia: a focus on drug and alcohol dependence. *HIV Med.* 2004; 5:30-33.

17. Liddicoat R, Horton NJ, Urban R, Maier E, Christiansen D, **Samet JH**. Assessing missed opportunities for HIV testing in medical settings. *J Gen Intern Med.* 2004; 19:349-356.

18. Palepu A, Horton NJ, Tibbetts N, **Meli S**, **Samet JH**. Uptake and adherence to highly active antiretroviral therapy among HIV-infected people with alcohol and other substance use problems: the impact of substance abuse treatment. *Addiction.* 2004;99:361-368.

19. Peek M, Han JH. Disparities in screening mammography current status, interventions, and implications [review]. *J Gen Intern Med.* 2004;19:184-194.

20. Peterson NB, **Friedman RH**, **Ash AS**, Franco S, **Carr PL**. Faculty self-reported experience with racial and ethnic discrimination in academic medicine. *J Gen Intern Med.* 2004;19:259-265.

21. Pope GC, Kautter J, Ellis RP, **Ash AS**, Ayanian JZ, Iezzoni LI, Ingber MJ, Levy JM, Robst, J. Risk adjustment of medicare capitation payments using the CMS-HCC model. *Health Care Financing Rev.* 2004; 25:119-141.

22. **Saitz R**, Helmuth ED, Aromaa SE, Guard A, Belanger M, Rosenbloom DL. Web-based screening and brief intervention for the spectrum of alcohol problems. *Prev Med.* 2004; 39:969-975.

23. **Saitz R**, **Horton NJ**, Larson MJ, Winter M, **Samet JH**. Primary medical care and reductions in addiction severity: a prospective cohort study. *Addiction.* 2004;100:70-78.

24. **Saitz R**, Larson MJ, **Horton NJ**, Winter M, **Samet JH**. Linkage with primary medical care in a prospective cohort of adults with addictions in inpatient detoxification: room for improvement. *Health Serv Res.* 2004;39:587-606.

25. **Samet JH**. Drug abuse and dependence. In: Goldman L and Auselio J, eds. *Cecil's Textbook of Medicine* 22nd ed. Philadelphia, Pa: WB Saunders; 2004:146-152.

26. **Samet JH**, Horton NJ, **Meli S**, **Freedberg KA**, Palepu A. Alcohol consumption and antiretroviral adherence among HIV-infected persons with alcohol problems. *Alc Clin Exp Res.* 2004;28:572-577.

27. **Samet JH**, Phillips S, **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.

28. Sweeney LP, **Samet JH**, Larson MJ, **Saitz R**. Establishment of a multidisciplinary Health Evaluation and Linkage to Primary Care (HELP) clinic in a detoxification unit. *J Addict Dis.* 2004;23:33-45.

29. Weedn VW, McDonald MD, Locke SE, Schreiber M, **Friedman RH**, Newell RG, Temoshok LR. Managing the community response to bioterrorist threats. *IEEE Eng in Med and Bio Magazine.* 2004;162-170.

30. Wines, JD Jr, **Saitz R**, **Horton NJ**, Travaglini-Lloyd C, **Samet JH**. Suicidal behavior, drug use and depressive symptoms after detoxification: a 2-year prospective study. *Drug Alcohol Depend.* 2004;76S:21-29.

Compensation and Advancement of Women in Academic Medicine: Is There Equity?

Arlene S. Ash, PhD; Phyllis L. Carr, MD; Richard Goldstein, PhD; and Robert H. Friedman, MD*

Background: **Women have been entering academic medicine in numbers at least equal to their male colleagues for several decades. Most studies have found that women do not advance in academic rank as fast as men and that their salaries are not as great. These studies, however, have typically not had the data to examine equity, that is, do women receive similar rewards for similar achievement?**

Objective: **To examine equity in promotion and salary for female versus male medical school faculty nationally.**

Design: **Mailed survey questionnaire.**

Setting: **24 randomly selected medical schools in the contiguous United States.**

Participants: **1814 full-time U.S. medical school faculty in 1995–1996, stratified by sex, specialty, and graduation cohort.**

Measurements: **Promotion and compensation of academic medical faculty.**

Results: **Among the 1814 faculty respondents (response rate, 60%), female faculty were less likely to be full professors than were men with similar professional roles and achievement. For example, 66% of men but only 47% of women ($P < 0.01$) with 15 to 19 years of seniority were full professors. Large deficits in rank for senior faculty women were confirmed in logistic models that accounted for a wide range of other professional characteristics and achievements, including total career publications, years of seniority, hours worked per week, department type, minority status, medical versus nonmedical final degree, and school. Similar multivariable modeling also confirmed gender inequity in compensation. Although base salaries of nonphysician faculty are gender comparable, female physician faculty have a noticeable deficit (−$11 691; $P = 0.01$). Furthermore, both physician and nonphysician women with greater seniority have larger salary deficits (−$485 per year of seniority; $P = 0.01$).**

Limitations: **This is a cross-sectional study of a longitudinal phenomenon. No data are available for faculty who are no longer working full-time in academic medicine, and all data are self-reported.**

Conclusions: **Female medical school faculty neither advance as rapidly nor are compensated as well as professionally similar male colleagues. Deficits for female physicians are greater than those for nonphysician female faculty, and for both physicians and nonphysicians, women's deficits are greater for faculty with more seniority.**

Ann Intern Med. 2004;141:205-212. www.annals.org

For author affiliations, see end of text.

See editorial comment on pp 238-240.

*Drs. Ash and Carr contributed equally to this manuscript.

Female medical school faculty have not advanced to senior academic ranks and positions in proportion to their numbers in academic medicine (1). Reports from many specialties and institutions have documented this situation (2–8). Certain specialties have actually reported a decline between 1995 and 2000 in the percentage of full professors who are women (for example, from 11% to 6% in emergency medicine and from 2% to 1% in orthopedic surgery). In 2000, only 8% of medical school chairs were women, and just 8 of 125 U.S. medical school deans were female (4 of them were interim) (9). Nonnemaker (10), using data on all U.S. medical schools and all U.S. medical school graduates from 1979 to 1993, found that women had continuing slower advancement to senior rank and that the proportion of female physicians entering academic medicine declined (10). However, that study had no job descriptors or measures of faculty performance and thus could not address the equity of these differences.

Female physicians also receive lower financial compensation, both in academic centers and in private practice (3, 8, 11); Baker, however, in examining salaries for young physicians in all settings (those with 2 to 9 years of experience in 1990) did not find the 41% greater salaries of men to be inequitable. In his model, differences in self-reported hours worked "explained" most of the observed salary difference, and differences in job characteristics, principally specialty and practice setting, accounted for the rest (11).

No large, detailed study in a nationally representative sample of institutions conducted across all medical school departments (including the basic sciences) has explored gender equity of faculty in advancement and compensation in academic medicine. Our study examines rich data from more than 1800 male and female academic faculty in all medical school departments at 24 randomly selected schools.

METHODS

Study Design

In 1995–1996, we conducted a national mailed survey (12) to examine the status of female, minority, and generalist academic medicine faculty. In the first stage of a 2-stage sampling plan, we sought 24 U.S. medical schools. Of the 126 medical schools listed by the Association of American Medical Colleges (AAMC) in 1995, we excluded 6 schools outside the contiguous United States because the AAMC considered them to be substantially different from the mainland schools. In addition, to obtain reasonable numbers of female and minority faculty from each institu-

tion, we excluded 14 schools that had fewer than 200 faculty, 50 female faculty, or 10 ethnic or racial minority faculty. Our 24 medical schools were randomly selected from the remaining 106 eligible medical schools. The resulting sample of schools was balanced across the AAMC's 4 regions of the United States and between public and private institutions.

In the second sampling stage, we selected full-time salaried faculty members from the 24 schools by using the 1995 AAMC Faculty Roster. The AAMC listed 17 434 faculty at the 24 schools; 720 faculty were excluded because they were in unique departments not found at other medical schools. Of the remaining 16 714 faculty, 4156 were women, 929 belonged to a racial or ethnic minority, and 869 were generalists. For each institution, sampling was stratified by the following: 4 areas of medical specialization (primary care, medical specialty, surgical specialty, and basic science), 3 graduation cohorts (received doctoral degree before 1970, between 1970 and 1980, and after 1980), and sex. We randomly sampled 6 faculty in each cell (school × medical specialty × graduation cohort × sex). The most senior graduation cohort cells were filled first. When a cell contained fewer than 6 people, we finished filling it with faculty who were from the same school, specialty, and sex but who were more junior. To obtain sufficient numbers of minority, generalist, and senior female faculty, we added all such faculty to the sample.

Data Collection and Survey Instrument

Our inclusion criteria required faculty to be full-time and currently employed at their AAMC-listed institution. We mailed 4405 surveys to sampled faculty, of which 1073 were ineligible because they had left their institution (n = 512), were not full-time (n = 510), had died (n = 11), or had participated in the pilot study (n = 9). The remainder (n = 31) were ineligible for other reasons. Nonrespondents among the eligible 3332 faculty received reminder postcards, follow-up telephone calls, and survey remailing, as necessary. Because of confidentiality concerns of the AAMC, we do not have further information on nonrespondents.

The self-administered questionnaire asked 177 questions about faculty demographic characteristics, current academic environment and support, academic productivity, rank, and faculty compensation. The survey was pretested by 45 medical school faculty at 3 institutions to ensure that respondents understood the meaning of the questions and could answer them appropriately. The Boston University School of Medicine Institutional Review Board approved the study.

Definitions of Analytic Variables

All reported data are from responses to survey questions. Career "seniority" was calculated as the number of years from first full-time faculty appointment (not necessarily at the current institution) until 1996. When the appointment year was missing, year of graduation from medical school + 4 was used in its place. For example, seniority equal to 25 years indicates either a first appointment in the 1970–1971 academic year or completion of schooling in 1967. We categorized race by using the AAMC classifications of white, majority, and 2 classes of minorities (13). Underrepresented minorities included black persons and most Hispanic persons, and nonunderrepresented minorities included Asian and Cuban persons. Missing race was imputed as white. Publications were specified as the career total number of any-authored articles in refereed journals; faculty who skipped this question were assigned a zero. To limit the influence of large outliers (for example, faculty reporting >500 publications or 120 hours of work per week), we coded publications in categories (0 to 9, 10 to 19, 20 to 39, 40 to 59, and ≥60) and top-coded "hours worked per week" at 80. "Chair or chief" is a marker for being a department chair or a division chief in 1996. "Physician" indicates faculty with a physician's degree (for example, MD or DO). Faculty responses to a request to divide 100% of their time into 4 categories (clinical, administrative, research, and teaching) yielded (continuous) "percent time in . . ." variables. Faculty with missing salaries (3%) were dropped from salary analyses, and those with missing rank (2.5%) were dropped from promotion studies.

We used the survey data to classify respondents' departments into the 4 prospectively identified types. "Primary care" includes general internal medicine and general pediatrics, family medicine, and geriatrics; "medical specialty" includes internal medicine and pediatric subspecialties, neurology, physical medicine, radiology, emergency medicine, anesthesia, and psychiatry; "surgical specialty" includes general surgery and its subspecialties, as well as obstetrics and gynecology; and "basic science" includes preclinical biological science.

Our outcomes were "salary" and "promotion." We used the term *salary* to refer to all pretax 1995–1996 academic-year faculty compensation, including clinical payments for the academic year (excluding fringe benefits, moonlighting, and consulting) and rounded to the nearest thousand. We defined the term *promotion* as having attained the rank of full professor by 1996.

Statistical Analysis

We used frequency distributions, means, and standard deviations to separately describe female and male respondents. We used linear regression to analyze salary and logistic regression to examine promotion. In each model, we adjusted standard errors using "school" as a clustering variable. The following additional predictors were used in both models: physician status, department type, minority status, chair or chief, school, seniority (either coded as a continuous variable or in 5-year categories to a maximum of ≥30), hours worked per week, and number of career publications (coded in categories [0 to 9, 10 to 19, 20 to 39, 40 to 59, ≥60]). In modeling compensation, we also adjusted

for current percentage of time in research and teaching. However, we did not use these time-allocation variables to predict promotion because we did not know their values historically. Because being a chair or chief is an outcome whose use as a predictor is controversial, we also examined the effect of dropping this predictor.

For each outcome, we constructed models based on both male and female faculty data and interpreted the size and statistical significance of sex-related coefficients of the model as measures of and tests for sex differences. This model directly answers questions such as "do women who are more senior experience larger deficits than less senior ones, and, if so, how much additional deficit per additional year of seniority?" We summarized the deficits in promotion for women within a faculty cohort such as "those with a first faculty appointment between 1975 and 1979" as follows. First, we suppressed the information as to which faculty were women and used the previously developed model to calculate, for each faculty, the probability of being a full professor (effectively assuming that all faculty were promoted "as if they were men"). We then compared the difference—the actual percentage of professors minus the predicted number—for women versus men in the cohort. If the women are underpromoted compared with men, their difference will be negative, whereas the men's difference will probably be close to, although not exactly, zero. Finally, we reported the risk-adjusted women's promotion deficit in the cohort as the women's difference minus the men's difference. We tested for the significance of this difference by using a 2-sample *t*-test. The risk-adjusted women's salary deficit was calculated and tested analogously.

Because we believed that salary structures might differ for physician and nonphysician faculty, as well as for male and female faculty and across department types (for example, basic science vs. surgery), we tested selected interactions among these variables for their potential importance as predictors. Specifically, we evaluated interactions of women by physician status, seniority, career publications, rank, and chair or chief, and interactions of physician status by seniority; underrepresented minority; department; career publications; rank; chair or chief; hours worked per week; and percentage of time in research, administration, and teaching. We retained such interactions when they were statistically significant at a *P* value less than 0.05. We believed that full professorship would rarely be attained in fewer than 10 years but that such promotion would be steadily attained during the subsequent 10 to 15 years for most persons who ever attain it. Thus, in our analysis, we restricted our modeling to faculty with 10 or more years of seniority, included an interaction between being female and seniority, and, to capture the expected leveling off, added a "long-term" marker for faculty with at least 25 years of seniority. Finally, because Baker's study (11) could be interpreted as finding that gender equity in promotion problems was solely a phenomenon of the prefeminist past, we tested a female × long-term interaction for its independent value in predicting full professor status. If, in fact, women and men hired since 1970 have been promoted comparably—even though the older cohort of female faculty did not fare so well—this interaction term would be significant, and its inclusion would cause the female × seniority interaction to lose its explanatory power. The test for whether sex affects the probability of being a full professor for faculty with at least 10 years of seniority is based on the significance of the indicator for female, whereas the test for a larger deficit for women of greater seniority is based on the joint significance of the female × seniority and female × long-term variables. We present the odds ratios, CIs, and *P* values for the resulting model in **Appendix Table 1** (available at www.annals.org).

Salary models were also used to test coefficients for their size and significance and to examine differences between expected and actual salaries for cohorts of women. The test for a gender difference in salary for nonphysician faculty in the first year is based on the significance of the female indicator; the test for a difference in salary for a female versus a male physician in the first year is based on the significance of the female physician interaction. The test for an increasing gender difference in salary is based on the significance of the female × seniority interaction. We report the salary model, its coefficients, CIs, and *P* values in **Appendix Table 2** (available at www.annals.org). We used Stata software, version 7 (Stata Corp., College Station, Texas), for all analyses.

Role of the Funding Sources

The Robert Wood Johnson Foundation funded the study but had no role in its design, conduct, or reporting or in the decision to submit the manuscript for publication.

RESULTS

Characteristics of the Faculty Sample

The 1814 respondents represent a response rate of approximately 60% for both male and female faculty (**Table 1**). Because of the stratified sampling, male and female respondents were similarly distributed by department, region of the country, and public–private status of their school. However, although we oversampled women who were senior faculty, female respondents were somewhat younger (mean age, 45 years vs. 47 years), more junior (only 31% of women vs. 38% of men had at least 15 years of career seniority), and less likely to be full professors (22% vs. 35%). Racial distributions were similar for men and women.

Advancement to Full Professorship

In unadjusted analyses, female faculty were less likely to be full professors than men of similar credentials (**Table 2**). For example, 66% of men with 15 to 19 years of seniority (that is, those first hired between 1976 and 1980) but only

Table 1. **Demographic and Professional Characteristics of Respondents***

Variable	Women (*n* = 873 [48%])	Men (*n* = 941 [52%])
Mean age ± SD, *y*	45 ± 9	47 ± 9
Race or ethnicity, *n (%)*		
White	715 (82)	753 (80)
Underrepresented minority	71 (8)	114 (12)
Nonunderrepresented minority	87 (10)	74 (8)
Physician status, *n (%)*	537 (61)	663 (70)
Region, *n (%)*		
Northeast	342 (39)	327 (35)
South	184 (21)	219 (24)
Midwest	159 (18)	197 (22)
West	183 (21)	182 (20)
Institution, *n (%)*		
Private	419 (48)	429 (46)
Public	450 (52)	504 (54)
Department category, *n (%)*		
Basic science	221 (26)	213 (23)
Medical specialty	164 (19)	159 (17)
Surgical specialty	142 (17)	176 (19)
Primary care	327 (38)	380 (40)
Career years of seniority, *n (%)*		
0–10	499 (57)	474 (50)
11–14	103 (12)	111 (12)
≥15	271 (31)	356 (38)
Career publications, *n (%)*		
0–9	392 (45)	332 (35)
10–19	144 (16)	119 (13)
20–39	163 (19)	186 (20)
40–59	90 (10)	120 (13)
≥60	84 (10)	184 (19)
Mean hours of work/wk ± SD	56 ± 11	58 ± 10
Mean time in research ± SD, %	29 ± 29	28 ± 29
Mean time in teaching ± SD, %	21 ± 15	19 ± 14
Chief or chair, *n (%)*	115 (13)	194 (21)
Rank, *n (%)*		
Full professor	190 (22)	322 (35)
Associate professor	226 (27)	237 (27)
Assistant professor	384 (45)	322 (35)
Instructor	48 (6)	40 (4)
Mean 1995 salary ± SD (in thousands), *$*	98 ± 45	125 ± 66

* Faculty with missing values of individual variables were dropped from percentage and mean calculations. Because sampling was stratified by school, sex, seniority, and department and was augmented to enhance numbers of senior faculty women and minority faculty, respondent percentages do not reflect national distributions of these characteristics for academic medical faculty.

47% of such women were full professors. **Table 2** also reveals that within each seniority cohort, female faculty were less likely than male faculty to have at least 40 publications; this finding emphasizes the importance of adjusting for such differences when examining equity. However, multivariable analysis also found substantial inequities in advancement for senior faculty women.

In the analysis for advancement (**Appendix Table 1**, available at www.annals.org), the only significant interaction was female × seniority (odds ratio, 0.90; $P = 0.003$), which suggests that each additional year of seniority was of substantially less value to women than to men in improving the chance of being a full professor. Underrepresented minority faculty were also less likely to have been promoted. In contrast, each of the following made full professorship more likely: being in a basic science department, having more career publications, being chair of a department or chief of an academic division, and working more hours. Three variations on this model yielded very similar odds ratio estimates (0.88 to 0.90) for the female × seniority interaction: 1) retaining only those 435 faculty with 10 to 20 years seniority; 2) adding a female × long-term marker (≥25 years), which was not significant ($P > 0.2$); and 3) dropping "chair or chief" as a predictor.

The cohort analysis (shown in the rightmost columns of **Table 2**) show the largest deficits in advancement for women among faculty hired before 1965 (44%), but notable deficits (22%) also persist for those hired as recently as 1970 to 1974.

Compensation

In the compensation analysis, the professional predictors of salary were having seniority (nearly $11 000 in annual compensation for each 10 years of seniority); having publications (for example, 40 publications added about $20 000 and 60 publications added >$30 000); being a physician (worth $43 000) and, especially, a physician in a medical or surgical specialty (worth, respectively, >$20 000 and >$50 000 more than the salary of nonphysicians in any department or physicians in primary care or basic sciences); being a chair or chief (worth $22 000); and working more hours (an 80-hour work week yielded almost $22 000 more than a 40-hour work week). In addition, each 10% of time spent in research was associated with a $3000 reduction in compensation for nonphysicians and a $7000 reduction for physicians, and each 10% of time spent in teaching (as opposed to clinical or administrative work) was associated with an almost $4000 reduction in compensation. Some differences were not based on profession: Female physicians received nearly $12 000 less than male physicians; women received almost $5000 less additional salary than men for each 10 years of seniority; and nonunderrepresented minority faculty received $7000 less than majority faculty. The model predicts, for example, that a white male primary care physician faculty with fewer than 10 publications will earn $96 214 in his first year; if he were a medical specialist, he would earn $116 003. A similarly situated female in either scenario will earn $11 691 less. With 10 years' seniority, the gender deficit increases by $4850 to $16 541. The female × seniority deficit shrinks (from $485 to $410 per year) in a sensitivity analysis that excluded faculty with 30 or more years of seniority (50 men and 25 women), and dropping "chair"

Table 2. **Attainment of Full Professor Rank and Publications by Sex and Seniority***

Seniority, *y*	Year of First Appointment	Faculty, *n*		Full Professor Rank, %		≥40 Publications, %		Risk-Adjusted Female Deficit†	*P* Value‡
		Women	Men	Women	Men	Women	Men		
0–4	1990–1995	241	220	1.2	3.6	1.2	5.4	NA	
5–9	1985–1989	219	211	4.6	4.3	6.4	14.2	NA	
10–14	1980–1984	129	148	20.2	24.3	23.5	35.3	4.8	>0.2
15–19	1975–1979	88	122	46.6	66.4	33.3	48.8	−5.7	>0.2
20–24	1970–1974	86	100	68.6	81.0	45.1	46.5	−22.4	<0.001
25–29	1965–1969	60	72	63.3	86.1	45.2	60.3	−21.7	<0.001
≥30	Pre–1965	25	48	52.0	93.8	36.0	66.7	−44.0	<0.001
Total		848	921	22.4	35.0	19.9	32.3	−11.0	<0.001

* NA = not available.
† Sex difference in observed minus expected percentage who are full professors (absolute percentage points). All expected differences are based on the probability of being a full professor among male faculty, as predicted by the model in Appendix Table 1 (available at www.annals.org).
‡ Each *P* value results from a *t*-test of H_0: women's (observed − expected) = men's (observed − expected) based on the total number of male and female faculty within the specified cohort.

as a predictor increases the magnitude of the female × seniority and female × physician estimated deficits (to −$568 and −$13 738, respectively).

The cohort analysis (**Table 3**) also suggests that the women's deficit is larger for women of more senior faculty rank, especially those hired before 1975. However, it also finds female salary deficits in every cohort, including a particularly large one among faculty hired since 1990. An analogous analysis found the salaries of female chairs and chiefs to be $17 800 less than those of male peers ($P < 0.001$).

Discussion

Our study confirms earlier findings that women in academic medicine have not reached senior academic ranks in proportion to their representation in medical school faculties. By considering and accounting for important professional characteristics (including number of career peer-reviewed publications) that independently affect faculty advancement, we have shown that women are significantly less likely to be full professors than comparably credentialed men. This is more than a pipeline phenomenon. Although ample numbers of women have entered academic medicine for at least the past 2 decades, the representation of women among full professors was only slightly higher in 1998 than in 1978 (10.5% vs. 7%) (14).

Our study found substantial deficits in academic rank for women, notably within cohorts whose first full-time appointment occurred between 1970 and 1985 and for women who became faculty before 1970. Although a few institutions have documented progress (15), most studies of women in academic medicine continue to find gender disparity in academic rank (1, 3, 16). Nonnemaker found female deficits in advancement in 15 consecutive national cohorts of academic faculty from 1979 to 1993; however, she did not have data on academic productivity, job characteristics, and performance to examine equity (10). Studies that have had such data have often been limited to 1 department or to 1 medical school (3–6, 15). Tesch and colleagues (16) conducted a national study and adjusted for important variables. In their study, 400 faculty hired in the 1980s from across the United States revealed issues with promotion similar to those found in our study, but they did not evaluate salary. We were able to examine faculty in all major medical school academic departments (basic science and clinical) and to account for important independent predictors of advancement, including numbers of peer-reviewed publications, hours worked per week,

Table 3. **Compensation by Sex and Seniority**

Seniority, *y*	Faculty, *n*		Year of First Appointment	Mean Salary ± SD (in thousands), $		Risk-Adjusted Female Deficit (in thousands)*, $	*P* Value†
	Women	Men		Women	Men		
0–4	241	217	1990–1995	85 ± 35	108 ± 61	−13.0	<0.001
5–9	217	204	1985–1989	94 ± 45	115 ± 61	−9.2	<0.001
10–14	126	147	1980–1984	105 ± 52	122 ± 62	−9.0	0.011
15–19	90	121	1975–1979	105 ± 48	134 ± 54	−10.4	0.005
20–24	88	99	1970–1974	113 ± 39	156 ± 78	−19.8	<0.001
≥25	86	123	Pre–1970	110 ± 51	140 ± 75	−24.0	<0.001
Total	848	911		98 ± 45	125 ± 66	−13.0	<0.001

* Sex difference in observed minus expected (thousands of) dollars in annual compensation. All expected differences are based on predicted salaries for male faculty by using the model in Appendix Table 2 (available at www.annals.org).
† Each *P* value results from a *t*-test of H_0: women's (observed − expected) = men's (observed − expected) based on the total number of male and female faculty within the specified cohort.

time spent in research and in teaching, and status as a chair or chief. Even after adjustment for these potential confounders, a concerning lack of equity in promotion to full professorship by sex remained.

Usual explanations for the scarcity of female full professors, other than simple discrimination, include women's lower motivation (17–19), their lack of mentorship (1, 20–23), sexual harassment (24–26), greater family responsibilities (20, 27, 28), less institutional support (6, 7), and the cumulative burden of many microinequities (29). However, most of these alternative explanations are not viable here. We found women to have similar motivation (30) and similar mentoring (31) as male faculty, and we did not find that gender bias or sexual harassment had noticeably affected academic productivity (12). Family responsibilities, however, did differentially weigh on female faculty, affect their academic productivity, and contribute to greater time to attaining senior rank (32). However, productivity differences do not fully explain the advancement deficit for women; at all levels of productivity, women are less likely to be full professors than are their male peers.

Equity in compensation has been equally problematic for female medical school faculty (3, 8), although 1 study (11) purported to show equity in compensation among recent graduates. However, as that cohort has aged, our study found increasing deficits. We found greater deficits for all female faculty with increasing years of seniority. For female physician faculty, in contrast to nonphysician faculty, we found a large additional deficit (−$11 691; $P = 0.01$). Others have found that the overall earnings differential between male and female physician faculty narrowed in the 1970s and 1980s (24% in 1972 [33], 19% in 1977 [34], and 15% in 1982 [35]) but did not disappear. Differences in specialty and practice style explain some of the salary deficits for women, as do differences in seniority, hours worked, and numbers of peer-reviewed publications (36). However, after adjustment for these and other faculty characteristics, female physician faculty are paid less than their male peers, and both physician and nonphysician female faculty experience greater deficits with greater seniority. We also found a significant salary deficit for non-underrepresented minority faculty.

We note that including a "chief or chair" indicator in models accepts women's lesser representation in leadership positions (13% vs. 21%) as a legitimate explanation for women's lower rank or salary. However, being passed over for a leadership position may be part of the same process that leads a woman to advance more slowly and be paid less than her male peers. The discrimination literature views variables that capture real differences in responsibility but may reflect discriminatory allocations as "tainted" (37). Dropping this variable had minimal effect on the promotion analysis but increased the estimated size of the salary deficit for women by about 17%. In addition, we found that female chairs and chiefs received $14 000 less than expected; this finding was based on the relations between professional characteristics and salary identified for their male peers.

Salary equity by sex or race is a legal as well as an ethical issue for employers. Pay discrepancies, typically associated with lower initial placement and slower promotions, have been found in successful gender discrimination lawsuits at universities (38). After accounting for the major professional factors that affect salary and advancement, substantial deficits for women and minorities remain; it is not obvious that additional legitimate factors, rather than discrimination, can account for these discrepancies. The gender deficits in both advancement and compensation are greater for women in more senior faculty positions.

Our findings in both advancement and salary parallel those of other studies in business, law, and academia (39). Starting salaries by sex for persons with an MBA, if experience is taken into account, tend to be approximately equal, but advancement for women is slower and salaries become increasingly disparate (39). In the profession of law, whether in private firms, corporations, or the judiciary, women are overrepresented in junior positions, are underrepresented in senior positions, and have lower salaries (39). Others have found the picture in academia to be the same and similar to our findings for academic medical faculty. The most recent female graduates start with salaries similar to those of their male colleagues, but by 3 to 8 years after a degree is earned, salary disparities appear and then increase with greater seniority. Gender differences in salary in science and engineering are greater than in the humanities. Overall, salary data for universities and colleges show almost no reduction of gender disparities between 1980 and 1996 (39).

The issues for women in science rather than medical academia are somewhat different; the "leaky pipeline" phenomenon is more potent here than in medical schools. Although nearly half (47%) of bachelor degrees in the sciences are awarded to women, only 38% of enrollees in graduate school in the sciences are women, and just 31% of PhDs in 1995 were awarded to women (40). In medicine, 40% of graduates are women, and, until recently, women have entered academia in higher proportion than their male colleagues (10). The cause of this decline in women entering medical academia over the past several years is unknown but could reflect resident and fellow awareness of the obstacles faced by female faculty.

Our work has limitations. Overall quality of academic performance is not fully captured by even our extensive data. However, there is no particular reason to believe that between 2 faculty of opposite sex with, for example, 50 publications each, legitimate "unmeasured factors," rather than gender-biased judgments, systematically favor the man. Although we have detailed data on many factors that may be associated with promotion, such as seniority and specialty, these data are self-reported; however, no evidence shows that any biases would be gender specific. Clearly,

number of peer-reviewed publications does not capture quality; however, promotion criteria are often (either formally or informally) linked to numerical "quotas." In addition, we have accounted for many other important factors that might affect rank, such as allocation of professional time, hours worked per week, seniority, and specialty.

Our information on compensation is also by self-report. However, we have no reason to suspect systematic differences by sex or other faculty descriptors in reporting professional income. One limitation is that significant differences in income can arise for faculty in the same surgical specialty if some faculty do fewer procedures. We have no data that speak to the existence, magnitude, or direction of such a difference by sex.

Our study is cross-sectional; thus, we know nothing about former faculty and whether men and women may have left in different numbers or for different reasons. Moreover, such data do not allow us to distinguish seniority from cohort effects. For example, we cannot say that the faculty hired since 1980, for whom no gender differences in promotion were apparent in 1996, are still subject to the same forces that led women hired in 1970 to experience a promotion deficit by 1996. A definitive answer to this question will not be available until after 2006. However, the fact that Baker (11) found gender equity in 1990 salaries for faculty hired within the preceding 10 years combined with our finding of salary inequities in 1996 for that cohort 6 years later suggests that gender inequities accrue over the course of a career. In the absence of corrective action, the gender inequities will probably continue to widen for current faculty as they become more senior. We do not know the level of academic productivity before promotion or productivity at the time of promotion, but we do know which faculty had not yet been promoted even though they had produced the number of publications reported.

Although 60% is a respectable response rate for a lengthy questionnaire administered to a nationally dispersed sample of academic physicians, nonresponders are sufficiently numerous that response bias could affect findings. Finally, our data are not as recent as we would wish (1995–1996); however, similarly rich, more recent data do not exist and data from the AAMC suggest that the gender gap in salaries persists. Thus, we believe that this study provides the best available data to address a very important issue.

Our work has many strengths. To our knowledge, it is the first study across all medical school departments, including clinical and basic science departments, in a national sample of medical schools to examine many key factors that affect academic advancement and compensation for men and women. Despite an adequate pipeline in academic medicine and sufficient years for women to achieve full professor rank, we found less advancement to full professor rank and lower salaries for women. Particularly in view of the decline in the numbers and proportion of women entering academic medicine (10), as well as the greater decrease in interest in an academic career for women compared with men during residency (41), medical schools should closely examine their environment for gender equity in promotion and compensation.

From Boston University School of Medicine, Boston, Massachusetts.

Acknowledgments: The authors thank Dr. Mark Moskowitz, deceased, for his thoughtful insights and contributions to this project. Thirty questions used in the survey are based on work by Drs. Linda Fried and Clair Francomano at the Johns Hopkins Medical Institutions.

Grant Support: By the Robert Wood Johnson Foundation, grant no. 019600.

Potential Financial Conflicts of Interest: None disclosed.

Requests for Single Reprints: Phyllis Carr, MD, Office of Student Affairs, Boston University School of Medicine, 715 Albany Street, L109, Boston, MA 02118; e-mail, plcarr@bu.edu.

Current author addresses are available at www.annals.org.

References

1. Promotion and tenure of women and minorities on medical school faculties. The American College of Physicians. Ann Intern Med. 1991;114:63-8. [PMID: 1983936]

2. **Tosi LL, Mankin HJ.** Ensuring the success of women in academic orthopaedics. Clin Orthop. 1998:254-63. [PMID: 9917692]

3. **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;335:1282-9. [PMID: 8857009]

4. **Wilkinson CJ, Linde HW.** Status of women in academic anesthesiology. Anesthesiology. 1986;64:496-500. [PMID: 3963456]

5. **Dial TH, Bickel J, Lewicki AM.** Sex differences in rank attainment among radiology and internal medicine faculty. Acad Med. 1989;64:198-202. [PMID: 2923647]

6. **Fried LP, Francomano CA, MacDonald SM, Wagner EM, Stokes EJ, Carbone KM, et al.** Career development for women in academic medicine: Multiple interventions in a department of medicine. JAMA. 1996;276:898-905. [PMID: 8782639]

7. **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;7:418-23. [PMID: 1506948]

8. **Carr PL, Friedman RH, Moskowitz MA, Kazis LE.** Comparing the status of women and men in academic medicine. Ann Intern Med. 1993;119:908-13. [PMID: 8215004]

9. **Bickel J, Clark V, Marshall Lawson R.** Women in U.S. academic statistics, 1999–2000. Washington, DC: Association of American Medical Colleges; 2000.

10. **Nonnemaker L.** Women physicians in academic medicine: new insights from cohort studies. N Engl J Med. 2000;342:399-405. [PMID: 10666431]

11. **Baker LC.** Differences in earnings between male and female physicians. N Engl J Med. 1996;334:960-4. [PMID: 8596598]

12. **Carr PL, Ash AS, Friedman RH, Szalacha L, Barnett RC, Palepu A, et al.** Faculty perceptions of gender discrimination and sexual harassment in academic medicine. Ann Intern Med. 2000;132:889-96. [PMID: 10836916]

13. **Petersdorf RG, Turner KS, Nickens HW, Ready T.** Minorities in medicine: past, present, and future. Acad Med. 1990;65:663-70. [PMID: 2102089]

14. AAMC faculty roster system: U.S. medical school faculty 1998. Washington, DC: Association of American Medical Colleges; 1998.

15. **Nickerson KG, Bennett NM, Estes D, Shea S.** The status of women at one academic medical center. Breaking through the glass ceiling. JAMA. 1990;264: 1813-7. [PMID: 2402039]

16. **Tesch BJ, Wood HM, Helwig AL, Nattinger AB.** Promotion of women physicians in academic medicine. Glass ceiling or sticky floor? JAMA. 1995;273: 1022-5. [PMID: 7897785]

17. **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;65:102-6. [PMID: 2302295]

18. **Beaty HN, Babbott D, Higgins EJ, Jolly P, Levey GS.** Research activities of faculty in academic departments of medicine. Ann Intern Med. 1986;104:90-7. [PMID: 3940508]

19. **Long JS, Allison PD, McGinnis R.** Rank advancement in academic careers: sex differences and the effects of productivity. American Sociology Review. 1993; 58:703-22.

20. **Graves PL, Thomas CB.** Correlates of midlife career achievement among women physicians. JAMA. 1985;254:781-7. [PMID: 4009912]

21. **Levinson W, Tolle SW, Lewis C.** Women in academic medicine. Combining career and family. N Engl J Med. 1989;321:1511-7. [PMID: 2811971]

22. **Ochberg RL, Barton GM, West AN.** Women physicians and their mentors. J Am Med Womens Assoc. 1989;44:123-6. [PMID: 2754173]

23. **Levinson W, Kaufman K, Clark B, Tolle SW.** Mentors and role models for women in academic medicine. West J Med. 1991;154:423-6. [PMID: 1877183]

24. **Komaromy M, Bindman AB, Haber RJ, Sande MA.** Sexual harassment in medical training. N Engl J Med. 1993;328:322-6. [PMID: 8419819]

25. **Grant L.** The gender climate of medical school: perspectives of women and men students. J Am Med Womens Assoc. 1988;43:109-10, 115-9. [PMID: 3411075]

26. **Baldwin DC Jr, Daugherty SR, Eckenfels EJ.** Student perceptions of mistreatment and harassment during medical school. A survey of ten United States schools. West J Med. 1991;155:140-5. [PMID: 1926843]

27. **Cole JR, Zuckerman H.** Marriage, motherhood and research performance in science. Sci Am. 1987;256:119-25. [PMID: 3810117]

28. **Neumayer L, Konishi G, L'Archeveque D, Choi R, Ferrario T, McGrath J, et al.** Female surgeons in the 1990s. Academic role models. Arch Surg. 1993;128: 669-72. [PMID: 8503771]

29. **Cole JR, Singer BA.** A theory of limited differences: explaining the productivity puzzle in science. In: Zuckerman H, Cole JR, Bruer RT, eds. The Outer Circle: Women in the Scientific Community. New York: W.W. Norton; 1991.

30. **Barnett RC, Carr P, Boisnier AD, Ash A, Friedman RH, Moskowitz MA, et al.** Relationships of gender and career motivation to medical faculty members' production of academic publications. Acad Med. 1998;73:180-6. [PMID: 9484191]

31. **Palepu A, Friedman RH, Barnett RC, Carr PL, Ash AS, Szalacha L, et al.** Junior faculty members' mentoring relationships and their professional development in U.S. medical schools. Acad Med. 1998;73:318-23. [PMID: 9526459]

32. **Carr PL, Ash AS, Friedman RH, Scaramucci A, Barnett RC, Szalacha L, et al.** Relation of family responsibilities and gender to the productivity and career satisfaction of medical faculty. Ann Intern Med. 1998;129:532-8. [PMID: 9758572]

33. **Kehrer BH.** Factors affecting the incomes of men and women physicians: an exploratory analysis. J Hum Resour. 1976;11:526-45. [PMID: 977939]

34. **Langwell KM.** Factors affecting the incomes of men and women physicians: further explorations. J Hum Resour. 1982;17:261-75. [PMID: 7130708]

35. **Ohsfeldt RL, Culler SD.** Differences in income between male and female physicians. J Health Econ. 1986;5:335-46. [PMID: 10282333]

36. **Dial TH, Grimes PE, Leibenluft E, Pincus HA.** Sex differences in psychiatrists' practice patterns and incomes. Am J Psychiatry. 1994;151:96-101. [PMID: 8267142]

37. **Paetzold RL, Willborn SL.** Tainted variables. In: The Statistics of Discrimination. Eagan, MN: West Group; 1999.

38. *Fish et al. vs. St. Cloud State University*. Described on the American Association of University Women Legal Advocacy Fund Web site: www.aauw.org/laf /cases/fish.cfm.

39. **Valian V.** Why So Slow: The Advancement of Women. Cambridge, MA: MIT Pr; 1998.

40. **Etzkowitz H, Kemelgor C, Uzzi B.** Athena Unbound: The Advancement of Women in Science and Technology. Cambridge, United Kingdom: Cambridge Univ Pr; 2000.

41. **Cain JM, Schulkin J, Parisi V, Power ML, Holzman GB, Williams S.** Effects of perceptions and mentorship on pursuing a career in academic medicine in obstetrics and gynecology. Acad Med. 2001;76:628-34. [PMID: 11401809]

Appendix Table 1. **Model To Predict Full Professor Status***

Variable	Odds Ratio (95% CI)	*P* Value
Sex		
Male	Reference†	
Female	1.28 (0.57–2.86)	>0.2
Seniority‡	1.35 (1.25–1.45)	<0.001
Female × seniority	0.90 (0.84–0.96)	0.003
Seniority of ≥25 y	0.19 (0.19–0.40)	<0.001
Race or ethnicity		
White	Reference†	
Underrepresented minority	0.42 (0.21–0.82)	0.01
Nonunderrepresented minority	0.61 (0.30–1.26)	0.18
Department category		
Primary care	Reference†	
Medical specialty	1.62 (0.95–2.75)	0.08
Surgical specialty	1.10 (0.61–1.99)	>0.2
Basic science	1.75 (1.21–2.53)	0.003
Career publications		
0–9	Reference†	
10–19	1.47 (0.69–3.13)	>0.2
20–39	2.74 (1.16–6.45)	0.02
40–59	13.40 (6.03–29.76)	<0.001
≥60	22.92 (9.36–56.12)	<0.001
Chair or chief	3.51 (2.31–5.33)	<0.001
Hours of work/wk§	1.39 (1.05–1.60)	<0.001

* The data pertain to all 482 male and 382 female respondents with at least 10 years of seniority.
† The reference group consists of white males who are in a primary care department, who have 0 to 9 publications, and who are not a chair or chief.
‡ Years beyond 10 since first full-time faculty appointment.
§ Hours of work/wk = each additional 10 hours over 40 worked/wk to a maximum of 80.

Current Author Addresses: Dr. Ash: General Internal Medicine, Boston University School of Medicine, 715 Albany Street, DOB 1108, Boston, MA 02118.
Dr. Carr: Office of Student Affairs, Boston University School of Medicine, 715 Albany Street, L109, Boston, MA 02118.
Dr. Goldstein: 37 Kirkwood Road, Brighton, MA 02135.
Dr. Friedman: MED Info Systems Unit, Boston University School of Medicine, 715 Albany Street, DOB 1102, Boston, MA 02118.

Appendix Table 2. **Model To Predict 1996 Compensation***

Variable	Compensation (95% CI), $	*P* Value
Intercept†	96 214 (81 160 to 111 269)	<0.001
Female	−332 (−8148 to 7473)	>0.2
Nonphysician	−43 131 (−55 493 to −30 768)	<0.001
Female × physician	−11 691 (−20 735 to −2647)	0.01
Seniority‡	1097 (642 to 1552)	<0.001
Female × seniority	−485 (−856 to −113)	0.01
Race or ethnicity		
White	Reference§	
Underrepresented minority	−1843 (−8990 to 5303)	>0.2
Nonunderrepresented minority	−6798 (−12 316 to −1280)	0.2
Department category		
Primary care	Reference§	
Basic science	1433 (−4618 to 7484)	>0.2
Medical specialty	3352 (−5868 to 12 572)	>0.2
Surgical specialty	7317 (−7692 to 22 326)	>0.2
Physician × department		
MD × medical specialty	19 789 (8096 to 31 482)	0.001
MD × surgical specialty	48 531 (27 182 to 69 881)	<0.001
Career publications		
0–9	Reference§	
10–19	9113 (4689 to 13 536)	<0.001
20–39	13 489 (6506 to 20 472)	0.001
40–59	20 466 (11 180 to 29 752)	<0.001
≥60	31 493 (24 078 to 38 908)	<0.001
Chair or chief	22 078 (14 879 to 29 277)	<0.001
Hours of work/wk‖	540 (308 to 773)	<0.001
Percentage of time in research¶	−297 (−463 to −131)	0.001
Physician × percentage of time in research	−375 (−563 to −188)	<0.001
Percentage of time in teaching¶	−375 (−568 to −183)	0.001

* Data pertain to all 848 female and 911 male respondents with no missing salary and other predictor information.
† The expected 1996 salary for a starting white male physician faculty member who is in primary care, who has <10 publications, who is neither a chair nor chief, and who works 40 h/wk (none of it in research or teaching). Expected salaries for other faculty are obtained by adding pertinent characteristics to $96 214.
‡ Years since first full-time faculty appointment.
§ The reference group consists of white men who are in a primary care department, have 0 to 9 publications, and are not a chair or chief.
‖ Hours of work/wk = each additional hour worked/wk beyond 40 to a maximum of 80.
¶ Percentage time = each 1% of time spent as indicated (vs. clinical and administrative activities).

Can We Use Automated Data to Assess Quality of Hypertension Care?

Ann M. Borzecki, MD, MPH; Ashley T. Wong, MA; Elaine C. Hickey, RN, MS; Arlene S. Ash, PhD; and Dan R. Berlowitz, MD, MPH

Objective: To determine whether extractable blood pressure (BP) information available in a computerized patient record system (CPRS) could be used to assess quality of hypertension care independently of clinicians' notes.

Study Design: Retrospective cohort study of a random sample of hypertensive patients from 10 Department of Veterans Affairs (VA) sites across the country.

Methods: We abstracted BPs from electronic clinicians' notes for all medical visits of 981 hypertensive patients in 1999. We compared these with BP measurements available in a separate vitals signs file in the CPRS. We also evaluated whether assessments of performance varied by source by using patients' last documented BP reading.

Results: When the vital signs file and notes were combined, a BP measurement was taken for 71% of 6097 medical visits; 60% had a BP measurement only in the vital signs file. Combining sources, 43% of patients had a BP reading of less than 140/90 mm Hg; by site this varied (34%-51%). Vital signs file data alone yielded similar findings; site rankings by rates of BP control changed minimally.

Conclusions: Current performance review programs collect clinical data from both clinicians' notes and automated sources as available. However, we found that notes contribute little information with respect to BP values beyond automated data alone. The VA's vital signs file is a prototypical automated data system that could make assessment of hypertension care more efficient in many settings.

(Am J Manag Care. 2004;10:473-479)

Obtaining valid data describing processes and outcomes of care is central to quality assessment and improvement. Traditionally, such data could be obtained from a variety of sources including administrative databases, medical records, and patient surveys. Administrative databases contain information typically collected for billing purposes or to track utilization, including demographics, diagnoses, and procedure codes. Such databases allow cost-efficient study of large numbers of cases but lack the clinically detailed information available from medical records.[1-3] Increasingly though, clinically detailed information such as laboratory and vital signs data are becoming incorporated into comprehensive information systems.[3,4] The completeness and accuracy of these data systems are often in question.[5,6] Consequently, assessment of their validity remains necessary.

Hypertension is an important condition whose treatment is in need of quality improvement.[7] It affects more than 50 million Americans and more than 1 million veterans.[8,9] Despite readily available, effective therapy for lowering blood pressure (BP) and preventing cardiovascular morbidity and mortality[8,10-12] most patients with hypertension have inadequate BP control.[13-18] In the 1999-2000 National Health and Nutrition Examination Survey, 69% of patients with a diagnosis of hypertension had a BP reading greater than or equal to 140/90 mm Hg.[8] Further, several studies have shown that despite reported familiarity and agreement with national hypertension guidelines, clinicians tolerate higher BPs than are recommended.[13,19-21]

Improving hypertension care requires ongoing assessment. Unadjusted BP control is the only widely used measure to assess hypertension care, used by both the Health Plan Employer Data and Information Set (HEDIS) and the Department of Veterans Affairs (VA) performance review program.[22,23] Unlike many other performance indicators that involve first examining automated data and then reviewing the medical chart if data are not available, hypertension assessment has traditionally relied solely on chart review.[24] This is at least in part because BP data may be recorded by several individuals. In most ambulatory clinics, a nurse takes an initial BP and documents this reading in an intake note. Clinicians may then do additional BP measurements, which are documented in their medical notes.

From the Department of Health Services, Boston University School of Public Health, Boston, Mass (AMB, ATW, DRB); the Center for Health Quality, Outcomes and Economic Research, Bedford VAMC, Bedford, Mass (AMB, ATW, ECH, DRB); and the Section of General Internal Medicine, Boston Medical Center, and Boston University School of Medicine, Boston, Mass (ASA, DRB).

This project was funded in part by the Department of Veterans Affairs Health Services Research and Development Service (grant SDR 99-300-1).

Address correspondence to: Ann M. Borzecki, MD, MPH, CHQOER, Bedford VAMC (152), 200 Springs Rd, Bedford, MA 01730. E-mail: amb@bu.edu.

For those settings with computerized patient record systems (CPRS), the initial vital signs information recorded by the nurse also may be entered into a separate data field of the record, which may then be readily extracted. In previous work, we found that only 1 BP reading was taken at most visits, which usually was present in the nurse's intake note.[14] In a setting where this information is entered directly into the computerized record, it is unknown how much information would be lost by examining only these automated vital signs data, and whether using only these data would impact quality measurement.

The current study compares the availability and agreement of BP measurements from an extractable data field of the CPRS with BP measurements obtained from clinicians' notes. We address the following 3 questions:

1. How complete is the automated BP information?
2. What factors are associated with discrepancies between the automated data and the clinicians' notes?
3. How do judgments about the quality of hypertension care vary based on the data sources used?

Lessons learned from our experience may be useful to other researchers and individuals interested in measuring healthcare quality.

METHODS

This is a retrospective cohort study that analyzed VA databases. The VA, as the largest integrated healthcare system in the United States, provides care to more than 4 million veterans and is considered to be a leader in establishing "a multifunctional integrated electronic medical record system."[25-27]

Study Subjects and Sites

We identified individuals with hypertension who were receiving regular outpatient medical care at 10 VA sites across the country during 1999. (A site comprises a hospital-based outpatient clinic and associated community-based outpatient clinics.) Selected sites had been entering BP measurements into a separate vital signs file of the CPRS and using electronic clinicians' notes for medical clinics, both as of at least January 1, 1999.

We used a national administrative VA database, the OutPatient Clinic file, to identify eligible subjects. To be eligible, patients needed to have at least 1 OutPatient Clinic–listed hypertension diagnosis (*International Classification of Diseases, Ninth Revision, Clinical Modification [ICD-9-CM]* code 401, 402 or 405) in 1998 and to be regular VA users (ie, ≥2 OutPatient Clinic–listed medical clinic visits at least 6 months apart in 1999). The study sample was randomly selected from among all eligible patients stratified by site. We sought 100 patients per site and achieved a final sample size of 981.

Data Collection and Sources

We used the VA's CPRS, known as the Veterans Health Information Systems and Technology Architecture, and the OutPatient Clinic file. The Veterans Health Information Systems and Technology Architecture, which is maintained at the hospital within a site, contains multiple files, including those with clinical data such as vital signs, laboratory and radiologic test results, pharmacy data, problem lists, and provider notes. It also contains an administrative-encounter file with diagnoses and procedures from all clinic visits that is transferred to a central VA data repository in Austin, Texas, and incorporated into the OutPatient Clinic file.[28] At the clinic level, BP measurements usually are taken by a nurse and either directly entered into a separate vital signs file with structured data entry fields in the CPRS, or reported on encounter forms and then entered by a clerk into this file. Additional BP data may be available through provider notes, which are either dictated and transcribed or typed directly into the provider notes file of the CPRS. (These clinical files are not yet routinely transferred centrally.)

Study data were collected during the 12 months in 1999. For automated data, patient demographics, *ICD-9-CM*–coded diagnoses, and medical clinic visit dates were obtained from the OutPatient Clinic file; BP measurements were extracted from the CPRS vital signs file. Vital signs file data were merged with OutPatient Clinic visit information such that a visit was assigned to each BP recording. For dates with multiple clinic visits, such as visits to primary care and a general surgical clinic, we assigned all BP recordings to the medical clinic.

Clinicians' notes from all medical visits of selected patients were obtained by accessing each site's local intranet and printing a hard copy. (As mentioned, the file containing these notes also is part of the CPRS, but the information would not be considered automated because it is free text.) An experienced nurse-abstractor then extracted note information including visit type, date, and BP. Blood pressure information from nurses' intake notes or clinicians' note entries that used an object template taken from the vital signs file were ignored because we were interested in whether the clinician took additional BP readings.

A 5% random sample of charts (a chart comprises all clinical notes on a given patient) was reviewed by one of

the authors (A.M.B.) for interrater reliability. Observed agreement on the presence and value of all readings was 96%. The only discrepancies found related to the presence of a BP reading rather than to its value. Such discrepancies were more likely to occur for patients who had more than 6 visits and more than 5 BP readings available, with one or the other reviewer missing an available BP.

Statistical Analyses

Completeness of Automated Blood Pressure Information. First, we determined whether BP measurements were available in the CPRS vital signs file. Our denominator consisted of all OutPatient Clinic–identified medical clinic visits in 1999. We examined the percentage of visits with at least 1 BP measurement, as well as the percentage with 2 or more measurements. If multiple identical BP values were found in the vital signs file for the same day, we deleted duplicates.

Next, we examined how much additional BP information would be obtained by combining vital signs file data and information available in clinicians' notes. We used the same denominator and eliminated duplicate values in the notes. We examined the amount of information lost by calculating the differences between sources overall and by site.

Discrepancies Between Sources. For visits with BP measurements available in both sources, we compared the number and value of BP recordings in the clinicians' notes with those in the vital signs file. We cross-tabulated visits by number of automated BP measurements against the number of BP measurements from the corresponding visit notes. Additionally, we checked whether each BP recording in the vital signs file had an exact match in the clinicians' notes, and noted the frequency of visits at which this match occurred. We checked for a match using both individual BPs from a given source and the average of available BPs.

We also examined whether BP documentation differences between sources (both in terms of BP presence and average BP value at a given visit) were because the BP was high and therefore the clinician was more likely to repeat the measurement and report it in his or her note. We used the average BP in the vital signs file at a given visit and determined whether it was high (≥140/90 mm Hg). We tested whether visits with vital signs file BPs versus those with BPs in both sources were more or less likely to have a high BP reading by using the chi-square test. We similarly compared the BPs of matching and nonmatching visits.

Variation in Judgment of Blood Pressure Control. We first examined differences in BP control (BP <140/90 mm Hg) at the individual visit level. We then determined patient-level control by calculating the average BP for each patient at his or her last visit of the year for which a BP value was available. We calculated the percentage of patients with a BP less than 140/90 mm Hg, examining results by source for the whole sample and testing for site differences by using the chi-square test. We again computed differences between combined sources and the vital signs file alone.

RESULTS

Baseline sample characteristics are presented in Table 1. The number of patients per site varied from 71 to 103 because of differential adoption of electronic notes by site. There were 6097 visits to primary care, medical subspecialty clinics, and urgent and emergent care. Of these, 3987 were primary care visits; 629 were subspecialty primary care visits (general internal medicine, geriatrics, women's clinic, hypertension, cardiology, spinal cord clinic); and 1481 were subspecialty, urgent care, or nursing visits.

How Complete Are the Data?

Sixty percent of all medical visits had at least 1 BP measurement in the vital signs file (Table 2). Combining automated (vital signs file) and clinician note information, 71% of visits had at least 1 recorded BP measurement (Table 2). Therefore, 11% of visits had a BP measurement in the clinicians' notes that was not

Table 1. Baseline Characteristics of 981 Patients With Hypertension

Characteristic	No. (%)
Age, y*	65.3 ± 11.1
Male sex	951 (97)
Nonwhite race	112 (11)
Number of antihypertensive medications	
0	87 (9)
1	249 (25)
2	309 (31)
3	194 (20)
≥4	142 (15)
Selected coexisting conditions	
Diabetes mellitus	322 (33)
Hyperlipidemia	442 (45)
Coronary artery disease	339 (35)
Cerebrovascular disease	79 (8)
Tobacco use	192 (20)

*Mean ± SD.

Table 2. Prevalence of Visits With at Least 1 or 2 Blood Pressure Measurements, by Source

	Mean ± SE, % (No.)	
Visits (n = 6097)	Vital Signs File	Vital Signs File + Notes*
With >1 blood pressure measurement	60.3 ± 0.6 (3677)	71.4 ± 0.6 (4350)
With ≥2 blood pressure measurements	1.9 ± 0.2 (115)	14.6 ± 0.5 (890)

*Notes are electronic clinicians' notes.

in the vital signs file. By site, the amount of available information lost using only the vital signs file varied from 1% to 35% ($P < .0001$; data not shown). Only 2% of visits had 2 or more BP measurements in the vital signs file, whereas 15% of visits had at least 2 BP measurements recorded by combined sources. Thus, 13% of visits had a second BP measurement in the clinicians' notes that was not in the vital signs file.

What Factors Are Associated With Discrepancies Between Sources?

A BP measurement was available in both the vital signs file and the clinicians' notes for 1361 visits. Seventy-nine percent (1070/1361) of these visits had only 1 BP measurement in each source. The BP measurement matched exactly for 50% (678/1361) of these visits. Of these matching visits, 99% (674/678) had only 1 BP measurement in each source.

Conversely, 50% (683/1361) of the BP measurements taken during visits did not match exactly. Fifty-eight percent (396/683) of these visits had only 1 BP measurement in each source. Of the 287 visits with more than 1 BP measurement available in either source, 188 visits had 1 BP value that matched and 237 visits had multiple BP measurements only in the clinicians' notes; 30 visits had multiple BPs only in the vital signs file.

Next we examined whether BP documentation differences between sources were related to BP level. Of the 4350 visits with a BP measurement in either source, 2316 had a BP measurement only in the vital signs file. Of these 2316 visits, 57% had a high BP reading (≥140/90 mm Hg), compared with 63% of visits with a BP measurement in both sources ($P = .02$). Thus, a clinician-noted BP value was more likely if the intake or vital signs file BP was high. Of the 1361 visits with a BP measurement in both sources, 78% of visits where the average BP values did not match had a measurement indicating uncontrolled BP in the vital signs file only, compared with 48% of visits where the BP measurements did match ($P < .0001$). This suggests the clinician was more likely to repeat the BP measurement when the intake value was high, resulting in nonmatching values, rather than just transcribe the intake BP (matching values).

How Do Judgments of Blood Pressure Control Vary Based on the Data Source?

At the visit level, only 61 visits would have been misclassified depending on the source. At 48 visits, the BP would have been classified as uncontrolled according to the vital signs file, but would have been considered controlled when the combined source was used. At 13 visits, the BP would have been classified as controlled according to the vital signs file, but would have been considered uncontrolled when the combined source was used.

At the patient level, using only automated data, the BP of 41% of the patients was controlled (see **Table 3**). Using both sources yielded similar results in terms of overall control and site rankings by percent control. Overall, 43% of patients had controlled BP, and the most a site changed ranking was by 2 places (Table 3). Thus, the extra information provided in the notes changed the assessment of BP control for a given patient in fewer than 2% of cases. In those few cases, the BP changed from uncontrolled to controlled.

If one assumed patients with missing BP measurements had uncontrolled BP (≥140/90 mm Hg) at their last visit, this assumption made minimal difference to overall results or site rankings, even for the site missing the most data (10/101 patients), when just automated information was used. Health Employer Data and Information Set and the VA's performance review program use the lowest available BP and assume the BP is uncontrolled if missing.[22,23] Analyzing by both these criteria made little difference to results (data not shown).

DISCUSSION

Current assessments of BP control rely largely on chart review and are therefore time-consuming and limited in scope. If valid BP data were available in automated form, this would make evaluations of BP control and quality of hypertension care more useful by encompassing more cases and allowing more timely feedback of information to providers, so that corrective actions would be more likely.[29]

In the present study, we found that most BP data were available in an automated form in the vital signs file of the VA's CPRS and that most medical visits had only 1 BP measurement available regardless of source. Of the 22% of visits with BP values available in both the automated data and the clinician's notes, half the time the BP in the clinician's note was a duplicate of the vital signs file BP, suggesting that the clinician was simply taking this information from the vital signs file or the nurse's note and incorporating it into his or her note. As expected, clinicians were more likely to repeat the BP measurement when the initial readings by nurses were high, but this situation did not occur very often. Most repeat measurements were not appreciably different, and their inclusion did not significantly affect judgments of control. Despite additional BP values available in the notes at 11% of visits, the percentage of patients with controlled BP did not change appreciably when comparing automated data with automated data plus notes.

Table 3. Blood Pressure Control by Source*

Site†	No. of Patients‡	% Patients With Blood Pressure <140/90 mm Hg		Rank Change	% Difference Between Sources§
		Vital Signs File	Vital Signs File + Notes		
1	103	50.5	50.5	0	0.0
2	101	46.5	48.5	-1	2.0
3	103	46.5	48.5	+1	2.0
4	100	40.4	43.0	-1	2.6
5	101	41.1	43.0	0	1.9
6	71	41.2	45.1	+2	3.9
7	101	38.5	39.6	-1	1.1
8	99	38.8	40.4	+1	1.6
9	101	36.6	38.6	0	2.0
10	101	31.7	33.7	0	2.0
Whole sample	**981**	**41.2**	**43.1**	—	**1.9**

*Blood pressure control was determined by using the average of blood pressure measurements at the last visit for which blood pressure was available.
†Sites are ranked 1 to 10, from highest to lowest percentage based on the vital signs file.
‡The total number of patients per site varied from 71 to 103. At site 6, electronic clinic notes were available for only 71 hypertensive patients.
§There were no significant differences between sites.

No other studies have attempted to validate automated BP readings or other vital signs data in this way. One other study by Goldstein et al examined recorded BP values and assessments of control, although its methods were somewhat different.[30] These researchers studied chart BPs, comparing the BP in the initial note by the nurse (which is comparable to the CPRS vital signs file) with BP measurements done by clinicians for 350 patients participating in a hypertension intervention study, at 2 separate primary care visits at the Palo Alto, California, VA, a site also used in our study. The BP was rechecked at 48% of visits where patients had uncontrolled BP and 38% of all visits. For approximately 25% of visits, patients who had an initial uncontrolled BP reading had a controlled BP at clinician recheck. Given their findings, Goldstein et al argue for including repeat BP measurements in quality assessments. However, our data show that, despite this site having the highest percentage (26%) of visits with 2 or more available BP measurements in the combined source, BP values at only 18 of 656 (2.7%) visits changed from uncontrolled to controlled when clinician information was considered (data not shown). The study by Goldstein et al was presented as a meeting abstract, so full details were not available. However, it is likely that dissimilar methods account for the discrepant results. Further, our methods better reflect those of current performance review programs.

Most of the available studies of automated data elements have used claims data to examine the validity of diagnoses or process measures.[24] Few have examined the use of automated clinical data for process or outcome measures, and only 1 other study has looked at BP. Kerr et al compared automated data from a central VA diabetes registry with medical record data (both electronic and paper) with respect to diabetes quality measures including the measurement and level of control of BP, low-density lipoprotein (LDL) cholesterol, and glycosylated hemoglobin (HbA_{1c}).[27] They also investigated whether combining information from both sources (compared with using either source alone) affected quality assessments for approximately 800 veterans receiving diabetes care in 1999. They found lower rates for all process measures using automated data, compared with either the medical record or with both sources combined. Unlike our study, they found fewer

BP measurements available in the automated data than the chart. For the process measure of the proportion of patients with a BP measurement in 1999, the respective proportions by source were 84%, 99%, and 99% for automated data, medical record, and combined sources. If we were to construct a similar measure, 98% of our sample's patients would have a BP measurement according to the automated data, whereas only 33% would have a BP measurement based on clinician note data alone; when the sources are combined, 100% would have a BP measurement. These differing results may have occurred because Kerr et al collected the BP data at different time periods within sources and perhaps because they used a different, less complete, automated source. However, like our study, the Kerr et al study found that overall rates for outcome measures, including the percentage of patients with BP less than 140/90 mm Hg, LDL cholesterol less than 130 mg/dL, or HbA_{1c} less than 9.5%, were comparable regardless of source, although they could not construct a combined BP control measure because BP data were measured at different time periods.

Thus, we are the first to examine an automated database with BP measurements and compare it with medical notes from the same time period. We found it yielded as much or more information than medical note review and gave assessments of performance comparable to those of a combined measure using automated data and notes.

Our data are already a few years old. However, increased automation and familiarity with the VA's CPRS have occurred over this time period. Blood pressure information is now more likely to be entered into the database and is more likely to be entered directly by the nurse who took the measure, as opposed to a third party. Thus, newer vital signs file data should be even more accurate and complete. (In this study, 673 visits had a BP value only in the notes, which is unlikely to occur in the present VA ambulatory clinic setting.)

There is no way to know the true reliability and accuracy of data entered into the vital signs file or clinicians' notes because we have no control over the measurement, documentation, and data entry process. This is true for all information systems and medical records.

Site 6 had the fewest patients because of difficulty finding patients with available electronic notes. Also at this site, assessment of performance regarding BP control varied the most by source. It had the lowest percentage of BP values entered into the vital signs file and the second highest percentage with BP values only in the notes. This site was clearly behind the others in the adoption of the electronic record and also in the entry of BP measurements into the vital signs file. This likely has changed with time so that there will be less of a discrepancy between the 2 sources.

Because we could not analyze by individual clinician, we do not know whether such assessments would vary by source. However, most visits were associated with only 1 BP value in either source, with more visits having information available in the vital signs file than in the notes. With increased adoption of the vital signs file, we would expect that those visits where only 1 BP value was available only in the notes would now have that information entered into the vital signs file. Therefore, clinician assessments should be consistent between sources.

All healthcare systems face the challenge of developing effective methods for assessing the quality of care. In the case of hypertension, such assessments require accurate BP information. Although this study only examined VA data systems and settings, it is likely that non-VA clinicians behave similarly with regard to BP measurement. We believe that by implementing or enhancing existing medical record systems with similar extractable data fields, other healthcare organizations also may find that they are able to make more efficient decisions about hypertension care. Moreover, such systems could incorporate other clinical data fields that could be likewise extractable, making clinically detailed information more readily obtainable and facilitating monitoring of various quality indicators across many medical conditions. These could include information such as whether a pneumococcal vaccination was given or a foot exam performed on a patient with diabetes.[27]

Current performance review programs collect clinical data from both clinicians' notes and automated sources as available.[22] Given the demonstrated completeness of automated BP data in the electronic record, we believe assessments of hypertension care can be made based on these data alone, making such evaluations more efficient. Where effective databases do not currently exist, the VA's vital signs file is a prototypical clinical computerized data system that could be easily adopted by other settings.

Acknowledgments

We thank Elaine Czarnowski, RN, for data abstraction and Marshall Goff for editorial assistance.

REFERENCES

1. **Fisher ES, Whaley FS, Krushat WM, et al.** The accuracy of Medicare's hospital claims data: progress has been made, but problems remain. *Am J Public Health.* 1992;82:243-248.

2. **Kashner TM.** Agreement between administrative data files and written medical records. *Med Care.* 1998;36:1324-1336.

3. **Iezzoni LI.** Assessing quality using administrative data. *Ann Intern Med.* 1997; 127:666-674.

4. **Jollis JG, Ancukiewicz M, DeLong ER, et al.** Discordance of databases designed for claims payment versus clinical information systems. *Ann Intern Med.* 1993; 119:844-850.

5. **Halpern J.** The measurement of quality of care in the Veterans Health Administration. *Med Care.* 1996;34 (3 suppl):MS55-MS68.

6. **Arts DGT, De Keizer NF, Scheffer GJ.** Defining and improving data quality in medical registries: a literature review, case study and generic framework. *J Am Med Inform Assoc.* 2002;9:600-611.

7. **Chobanian AV.** Control of hypertension—an important national priority. *N Engl J Med.* 2001;345(7):534-535.

8. **Hajjar I, Kotchen TA.** Trends in prevalence, awareness, treatment, and control of hypertension in the United States, 1988-2000. *JAMA.* 2003;290:199-206.

9. **Yu W, Ravelo AL, Wagner TH, et al.** The cost of common chronic diseases in the VA health care system. Presented as an abstract at: 20th Annual Meeting of the Department of Veterans Affairs Health Services Research and Development Service; February 13-15, 2002; Washington, DC.

10. **SHEP Cooperative Research Group.** Prevention of stroke by antihypertensive drug treatment in older persons with isolated systolic hypertension: final results of the Systolic Hypertension in the Elderly Program (SHEP). *JAMA.* 1991;265:3255-3264.

11. **Amery A, Birkenhager W, Brixko P, et al.** Mortality and morbidity results from the European Working Party on High Blood Pressure in the Elderly trial. *Lancet.* 1985;1:1349-1354.

12. **Collins R, Peto R, MacMahon S, et al.** Blood pressure, stroke, and coronary heart disease, part 2: short-term reductions in blood pressure: overview of randomized drug trials in their epidemiological context. *Lancet.* 1990;335:827-838.

13. **Berlowitz DR, Ash AS, Hickey EC, et al.** Inadequate management of blood pressure in a hypertensive population. *N Engl J Med.* 1998;339:1957-1962.

14. **Borzecki AM, Wong AT, Hickey EC, et al.** Hypertension control: how well are we doing? *Arch Intern Med.* 2003;163:2705-2711.

15. **Frijling BD, Spies TH, Lobo CM, et al.** Blood pressure control in treated hypertensive patients: clinical performance of general practitioners. *Br J Gen Pract.* 2001;51:9-14.

16. **Joffres MR, Hamet P, Rabkin SW, et al.** Prevalence, control and awareness of high blood pressure among Canadian adults. *CMAJ.* 1992;146:1997-2005.

17. **Colhoun HM, Dong W, Poulter NR.** Blood pressure screening, management and control in England: results from the health survey for England 1994. *J Hypertens.* 1998;16:747-752.

18. **Meissner I, Whisnant JP, Sheps SG, et al.** Detection and control of high blood pressure in the community. Do we need a wake-up call? *Hypertension.* 1999;34: 466-471.

19. **Oliveria SA, Lapuerta P, Mccarthy BD, et al.** Physician-related barriers to the effective management of uncontrolled hypertension. *Arch Intern Med.* 2002; 162:413-420.

20. **Hyman DJ, Pavlik VN.** Self-reported hypertension treatment practices among primary care physicians. *Arch Intern Med.* 2000;160:2281-2286.

21. **Hajjar I, Miller K, Hirth V.** Age-related bias in the management of hypertension: a national survey of physicians' opinions on hypertension in elderly adults. *J Gerontol Med Sci.* 2002;57A:M487-M491.

22. **Sennett C.** Implementing the new HEDIS hypertension performance measure. *Manag Care.* 2000;9(4 suppl):2-17.

23. **Veterans Health Administration/Department of Defense.** *FY 2002 VHA Performance Measurement System Technical Manual.* Washington, DC: The Office of Performance and Quality, Veterans Health Administration; March 2002. Available at: http://www.oqp.med.va.gov/cpg/ HTN/P/HTN_3_8_02_techman.doc. Accessed December 20, 2003.

24. **Dresser MVB, Feingold L, Rosenbranz SL, Coltin KL.** Clinical quality measurement: comparing chart review and automated methodologies. *Med Care.* 1997;35:539-552.

25. **Ashton CM, Septimus J, Petersen NJ, et al.** Healthcare use by veterans treated for diabetes mellitus in the Veterans Affairs medical care system. *Am J Manag Care.* 2003;9:145-150.

26. **Kizer KW.** The "new VA": a national laboratory for health care quality management. *Am J Med Qual.* 1999;14:3-20.

27. **Kerr EA, Smith DM, Hogan MM, et al.** Comparing clinical automated, medical record, and hybrid data sources for diabetes quality measures. *Jt Comm J Qual Improv.* 2002;28:555-565.

28. **VA Information Resource Center.** Toolkit for new users of national VA data at the Austin Automation Center. Available at: http://www.virec.research.med.va.gov/Support/Training-NewUsersToolkit/Toolkit.htm Accessed December 20, 2003.

29. **Mugford M, Banfield P, O'Hanlon M.** Effects of feedback of information on clinical practice: a review. *BMJ.* 1991; 303:398-402.

30. **Goldstein MK, Hoffman BB, Coleman RW, et al.** Differences between initial and repeat blood pressure measurements during the same clinic visit in primary care practice. Presented as an abstract at: 19th Annual Meeting of the Department of Veterans Affairs Health Services Research and Development Service; February 14-16, 2001; Washington, DC.

In A Page
OB/GYN & Women's Health

Phyllis L. Carr, MD
Associate Dean of Students
Associate Professor of Medicine
Boston University School of Medicine
Boston, Massachusetts

Hope A. Ricciotti, MD
Assistant Professor, Obstetrics, Gynecology, and Reproductive Biology
OB/GYN Clerkship Director
Women's Health Education Theme Director
Harvard Medical School
Beth Israel Deaconess Medical Center
Boston, Massachusetts

Karen M. Freund, MD, MPH
Professor of Medicine
Boston University School of Medicine
Director, Center of Excellence in Women's Health
Boston University School of Medicine
Boston, Massachusetts

Scott Kahan, MD
Intern, Franklin Square Hospital
Baltimore, Maryland

Blackwell Publishing

Blackwell Publishing, Inc., 350 Main Street, Malden, Massachusetts 02148-5018, USA
Blackwell Publishing Ltd, 9600 Garsington Road, Oxford OX4 2DQ, UK
Blackwell Science Asia Pty Ltd, 550 Swanston Street, Carlton, Victoria 3053, Australia

03 04 05 06 5 4 3 2 1

ISBN: 1-4051-0380-9

Library of Congress Cataloging-in-Publication Data

In a page OB/GYN & women's health / [edited by] Phyllis L. Carr ... [et al.].
p. cm.
Includes index.
ISBN 1-4051-0380-9 (alk. paper)
1. Obstetrics—Handbooks, manuals, etc. 2. Gynecology—Handbooks, manuals, etc.
3. Women's health services—Handbooks, manuals, etc.
I. Title: OB/GYN & women's health. II. Carr, Phyllis L.

RG103.I5 2003
618—dc22

2003045148

A catalogue record for this title is available from the British Library

Acquisitions: Beverly Copland
Development: Kate Heinle
Production: Debra Lally
Cover design: Gary Ragaglia
Interior design: Visual Perspectives
Typesetter: TechBooks in Delhi, India
Printed and bound by Sheridan Books in Ann Arbor, MI

For further information on Blackwell Publishing, visit our website:
www.blackwellpublishing.com

Medical Student Review

Are you overwhelmed by the sheer mass of information that you need to know? Now you can find all of the information you need about a disease, all on one page! *In A Page OB/GYN & Women's Health* is your timesaving solution to long hours and busy wards.

In this innovative one-page format you can review topics such as etiology/pathophysiology, differential diagnosis, presentation, diagnosis/evaluation, treatment, and prognosis/complications options for 126 topics.

- Organized by a woman's life span—beginning in adolescence and ending with issues of elderly women
- Integrated approach to Women's Health, Obstetrics, and Gynecology
- Written by students, residents, practitioners, and reviewed by specialists, ensuring you have the most up-to-date, accurate information
- Emphasizes high-yield information and the most commonly encountered diseases, disorders, and medical issues
- Provides you with a visually distinct and unique way to help you learn and remember
- Fits in your white coat pocket for on-the-spot reference

Praise for In A Page

"I think this format is extremely useful for both upper class medical students as well as interns and residents. You can review long detailed discussions in a matter of minutes."

Sean Armin, 3rd year medical student, UCLA Medical School

"It is easy to go to the section you are looking for and the bullet format is easy to read and follow. This book provides a good overview of the pertinent information needed when you see a patient with one of these problems."

Christina Thiruvathukal, 4th year medical student, George Washington University

"This quick summary of the essential facts makes for a quick reference text ideal for the first look at a disease and for rapid revision before ward rounds and exams. The pertinent facts are distilled into the minimum of text without obscure or elaborate details."

Sally-Ann Price, University of Leeds, UK, 5th year

ISBN 1-4051-0380-9

9 781405 103800

Blackwell Publishing

www.blackwellpublishing.com

Monitoring Depression Care

In Search of an Accurate Quality Indicator

Andrea Charbonneau MD, MSc, Amy K. Rosen, PhD,†‖ Richard R. Owen, MD,‡ Avron Spiro III, PhD,**††¶ Arlene S. Ash, PhD,§ Donald R. Miller, ScD,†‖ Lewis Kazis, ScD,†‖ Boris Kader, PhD,† Fran Cunningham, PharmD,†† and Dan R. Berlowitz MD, MPH†‖*

Background: Linking process and outcomes is critical to accurately estimating healthcare quality and quantifying its benefits.

Objectives: The objective of this study was to explore the relationship of guideline-based depression process measures with subsequent overall and psychiatric hospitalizations.

Research Design: This is a retrospective cohort study during which we used administrative and centralized pharmacy records for sample identification, derivation of guideline-based process measures (antidepressant dosage and duration adequacy), and subsequent hospitalization ascertainment. Depression care was measured from June 1, 1999, through August 31, 1999. We used multivariable regression to evaluate the link between depression care and subsequent overall and psychiatric hospitalization, adjusting for patient age, race, sex, socioeconomic status, comorbid illness, and hospitalization in the prior 12 months.

Subjects: We studied a total of 12,678 patients from 14 Northeastern VHA hospitals.

Results: We identified adequate antidepressant dosage in 90% and adequate duration in 45%. Those with adequate duration of antidepressants were less likely to be hospitalized in the subsequent 12 months than those with inadequate duration (odds ratio [OR], .90; 95% confidence interval [CI], .81–1.00). Those with adequate duration of antidepressants were less likely to have a psychiatric hospitalization in the subsequent 12 months than those with inadequate duration (OR, .82; 95% CI, .69–.96). We did not demonstrate a significant link between dosage adequacy and subsequent overall or psychiatric hospitalization.

Conclusions: Guideline-based depression process measures derived from centralized data sources offer an important method of depression care surveillance. Their accuracy in capturing depression care quality is supported by their link to healthcare utilization. Further work is needed to assess the effect of implementing these quality indicators on depression care.

Key Words: depression, quality of care, linking process and outcomes

(*Med Care* 2004;42: 522–531)

Depression is a pernicious and prevalent illness associated with increased morbidity, mortality, healthcare costs, utilization, and disability.[1–3] Depression is a chronic, often relapsing illness, especially if undertreated initially.[2,4] Most patients who seek care for depression are treated exclusively in primary care settings where at least 50% of patients are undiagnosed and 40% to 55% are insufficiently treated.[5] Clinical guideline development (1993) evolved in response to these well-documented inadequacies in the recognition and treatment of depression,[6–9] and a plethora of investigations have recently outlined efficacious quality improvement strategies for advancing depression care.[10–13] On review of this evidence, the U.S. Preventive Services Task Force recently endorsed routine screening for depressive disorders in primary care settings with established systems for diagnosis, effective treatment, and follow up.[14,15]

The national imperative to improve the quality of health care[16] has included initiatives to translate depression quality improvement strategies into clinical practice. Aligning depression quality improvement with methods used in managing other chronic illnesses has been an important step for depression care.[17] Depression management systems have demonstrated improved short- and long-term outcomes of

From the *Division of General and Geriatric Medicine, University of Kansas School of Medicine; the †Center for Health Quality, Outcomes, and Economic Research, Edith Nourse Rogers Memorial Veterans Hospital, Bedford, Massachusetts; the ‡Center for Mental Healthcare and Outcomes Research, Central Arkansas Veterans Healthcare System, Little Rock, Arkansas; the §Section of General Internal Medicine, Boston Medical Center, Boston, Massachusetts; the Departments of ‖Health Services and ¶Epidemiology, Boston University School of Public Health, Boston, Massachusetts; **Boston Veterans Affairs Healthcare System, Boston, Massachusetts; and ††Pharmacy Benefits Management, Edward Hines, Jr. VA Hospital, Hines, Illinois.

This work was supported by the Department of Veterans Affairs, Health Services Research and Development Service, Grant # CPI 99-134, and MNH 98-001 (Mental Health QUERI).

Reprints: Andrea Charbonneau, MD, MSc, Division of General and Geriatric Medicine, University of Kansas Medical Center, 3901 Rainbow Blvd., Wescoe 5026, Kansas City, KS 66160. E-mail acharbonneau@kumc.edu.

ISSN: 0025-7079/04/4206-0522
DOI: 10.1097/01.mlr.0000127999.89246.a6

depression severity and persistence, employment retention, functional status, and patient satisfaction.[18,19] Although barriers to implementing similar programs in naturalistic clinical settings exist,[20] the imperative to improve and protect healthcare quality in this era of health system reorganization mandate persistence in overcoming them.

Successful quality improvement depends on accurate information about healthcare quality; seminal health services literature directs us to health system structure (number of area hospitals, clinics, providers), process (health care rendered to patients), and outcomes (eventual consequences of health care) for this information.[21,22] Observing expected relationships between the 3 is especially powerful; ie, appropriate process of care results in a positive healthcare outcome. Few observational studies have established this link between process and outcomes for depression[23–25]; this is important not only for establishing accuracy, or internal validity, of a process measure, but also for quantifying the benefits of appropriate care. It could also lend clinical credibility to the measure (ie, clinicians could be more likely to comply with clinical care profiling that uses process measures clearly linked to improved outcomes).

Healthcare utilization is a fundamental general healthcare outcome. Depression effects increased health services utilization,[2,26,27] and recent work has identified it as an independent predictor of inpatient hospital readmission after adjusting for age, comorbidities, and functional status.[28] Establishing a link between depression process measures and healthcare utilization would lend substantial weight to the predictive validity of these process measures.[29]

We have previously used centralized administrative and pharmacy records of the Veterans Health Administration (VHA) to define guideline-based depression process measures, to assess depression care using these measures, and to identify patient- and provider-level predictors of adequate depression care.[30] We now explore the relationship of these process measures of depression care with subsequent overall and psychiatric hospitalizations.

METHODS

Study Design and Sample Definition

We conducted a retrospective cohort study of patients from 14 VHA medical centers in New England and upstate New York, relying entirely on existing databases. Full details are reported elsewhere.[30] In brief, we used VHA centralized administrative and pharmacy records to define a depressed cohort that underwent antidepressant treatment during a 3-month period in 1999. Subject eligibility criteria were as follows: at least 1 *International Classification of Diseases*, 9th edition, Clinical Modification (ICD-9-CM) diagnosis code 296.2x or 296.3x (major depression single or recurrent episode, respectively) noted from a psychiatry, primary care, emergency, or social work clinical setting during October 1, 1997, through September 30, 1999 (FY'98 or FY'99) or at least 1 ICD-9-CM diagnosis code 311.xx (depression not otherwise specified [NOS]) in a primary care clinical setting during FY'98 or FY'99, exclusive of other depression diagnosis codes; and receipt of at least 2 antidepressant from a VHA pharmacy during the time period of depression care profiling (June 1, 1999, through August 31, 1999).

We excluded patients with comorbid schizophrenia and/or bipolar disorder, because these illnesses can produce symptoms that could be confused with depression, or if combined with depression, will often be best treated by addressing the underlying chronic mental illness. Comorbid schizophrenia and/or bipolar disorder were noted by 2 outpatient or 1 inpatient ICD-9-CM diagnosis code 295.00–295.95 (schizophrenia) or 296.00–296.89, excluding 296.20–296.36 (bipolar disorder) during FY'98 or FY'99.

We identified 27,665 patients with depression during FY'98 and FY'99, representing 8.5% of the patient population in our 14 study sites. We eliminated 4809 (17%) because of comorbid schizophrenia and/or bipolar disease. Of those remaining, 19,119 (84%) were treated with antidepressant medication during FY'99. Of these, 12,678 received antidepressants during June 1, 1999, through August 31, 1999, referred to henceforth as the 3-month profiling period; this represents the final sample for which the process of depression care can be linked to subsequent hospitalization. Their characteristics are reported in Table 1.

Depression Stage

We stratified patients by depression stage because there are important stage-specific recommendations in the depression guidelines that might affect care. We identified depression stage using clean periods, defined as the period of time before the initial visit for depression that was free of both depression-coded clinical visits and antidepressant prescriptions. We applied a 6-month clean period in defining subject depression stage: acute (first 3 months after initial diagnosis), continuation (4–9 months), or maintenance (greater than 1 year after initial diagnosis).[31,32]

Independent Variables

Guideline-Based Depression Process Measures

We compared 3 dimensions of antidepressant therapy with clinical guideline benchmarks using the 1997 VHA Depression Guidelines,[7] a compilation of recommendations from the Agency for Healthcare Research and Quality (AHRQ, formerly AHCPR) and the American Psychiatric Association depression guidelines.[6,8] *Dosage adequacy* was achieved when the antidepressant average daily dosage during the 3-month profiling period met the guideline-recommended minimum daily dosage. It was calculated in the following manner: number of prescribed tablets × strength

TABLE 1. Sample Characteristics

Characteristic	Summary Statistics (N = 12,678)
Age	
Mean ± SD years	57 ± 14
Sex (%)	
Female	8.1
Race (%)	
Black	5.0
White	78.2
Other	1.7
Missing	15.1
Marital status (%)	
Not married	51
Missing	1
Annual income	
Mean ± SD (in $1000)	37 ± 24
Missing (%)	1
Comorbidity index	
Mean ± SD	2.5 ± 2.1
Anxiety disorder (%)	40
PTSD (%)	36
Alcoholism (%)	27
Outcomes	
Dosage adequacy	
Yes (%)	90
Duration adequacy	
Yes (%)	45
Follow-up visit adequacy, acute stage, Yes (%)	62

SD = standard deviation; PTSD = posttraumatic stress disorder.

(mg) tablet/number of prescription days. This average strength per day was compared with the guideline-recommended minimum daily dosages and defined as adequate if these were met, resulting in a dichotomous outcome variable.

Any of 22 antidepressants (Appendix A) prescribed during the 3-month profiling period were analyzed; however, antidepressant exclusion criteria were specified to most accurately capture the process of care. These included: 1) concomitant (≥7 days overlapping) prescriptions for trazodone, bupropion, or any tricyclic antidepressant were not analyzed for dosage or duration adequacy because these medications could be used for therapeutic indications other than depression (ie, smoking cessation, insomnia); and 2) if patients received multiple antidepressants that did not have the same dosage adequacy, they were assigned the dosage adequacy of the antidepressant prescribed for longer.

The appropriateness of adjunctive medications occasionally used to treat depression such as benzodiazepines, valproate, lithium, or antipsychotic regimens was not addressed in this study as a result of nonspecific guideline recommendations for their use.

Duration adequacy for each patient characterized the overall length of therapy with any eligible antidepressant during the profiling period. It was calculated as follows for continuation and maintenance-stage patients: number of days without antidepressant medication/number of days in the 3-month period. We recognized that for continuation and maintenance-stage patients, inadequate duration of therapy might actually represent cessation of therapy secondary to cure or patient self-referral to private-sector care. Therefore, we conducted a sensitivity analysis excluding continuation and maintenance-stage patients without antidepressants during the 6 months after the profiling period (896 patients from 11,900); this demonstrated similar conclusions (49% with adequate duration of antidepressants vs. 45%). Therefore, we hold all stages accountable for continuous dosing during the profiling period. For acute-stage patients, the calculation of duration adequacy was as follows: number of days without antidepressant medication/number of days in 3-month period after the first prescription.

Duration adequacy was defined as a dichotomy, with inadequate duration being >21% of the profiling period without antidepressants. This somewhat arbitrary boundary (.21) translates into 3 weeks of the 3-month period (or 1 week per month) and is consistent with other definitions of continuous dosing in the literature.[31] Sensitivity analyses of the duration adequacy definition were performed using different boundaries, as well as using duration adequacy as a continuous variable. Our findings were not affected by these variations.

Follow-up visit adequacy was determined only for the acute stage because explicit guideline-based follow-up recommendations exist only for this group. Adequacy was defined as at least 3 visits to primary care or psychiatry clinics within 3 months of the initial depression encounter. At least 2 visits in addition to the initial one within the first 3 months of diagnosis accords with guideline-recommended follow-up care.

Defining Potential Confounders of Relationship Between Depression Care Process and Outcomes

We adjusted for several potential confounders of the link to subsequent hospitalization.[28,33,34] Patient age, sex, race, and marital status were identified from the administrative files. We used mean income per zip code from the 1998 Internal Revenue Service tax returns file as a proxy for individual income as a result of uncertainty of the accuracy of reported income in our administrative data. We determined prior hospitalization (Y/N) during the 12 months before hospitalization outcome assessments.

We supplemented missing patient sociodemographic data (excluding income) with the 1999 Large Health Survey of Veteran Enrollees,[35] which collected the Medical Outcomes Study SF-36 adapted for use in veterans. This complemented 5% of missing race data, mostly by identifying white race.

We identified comorbidity burden using an unweighted count of up to 30 medical conditions, derived from the General Health Questionnaire of the Medical Outcomes Study,[36] and used in the Veterans Health Study to assess the relationship of comorbidities to health-related quality of life.[37] We demonstrated similar conclusions by using the Charlson index (data not shown). Comorbid psychiatric conditions, including anxiety disorder, alcoholism, and posttraumatic stress disorder (PTSD), were coded individually because their impact on the process of depression care could be greater than other comorbid medical conditions.[5,38] All comorbidities were identified from inpatient and outpatient VHA administrative data during FY'98 and FY'99.

Dependent Variables

Subsequent Overall and Psychiatric Hospitalizations

We identified subsequent inpatient hospitalizations from the Patient Treatment File, which documents all VHA inpatient hospital stays. Observational (overnight), acute care, and new extended care hospitalizations during the 12 months after the depression care profiling period were identified for the utilization subsample. We differentiated subsequent psychiatric hospitalization from all others using ICD-9-CM diagnoses responsible for the length of stay: 291.xx–292.xx, 295.xx–298.xx, 300.xx–311.xx. A sensitivity analysis exclusively using principal diagnosis (ie, primary reason for admission) did not reveal significant differences. We defined subsequent hospitalization as a dichotomy (any vs. none) in all regression analyses.

Analysis

We identified the number and type of hospitalizations during the 12 months after the depression care profiling period. We used logistic regression to examine the relationship of depression care adequacy with subsequent hospitalization, adjusting for patient age, sex, race, income, marital status, hospitalization in the prior year, and comorbid illness burden. Because depression care process could affect subsequent psychiatric hospitalization differently from other hospitalization types, we evaluated its relationship with overall hospitalizations and psychiatric hospitalizations in separate regressions.

We were concerned that different combinations of depression care adequacy might have modified outcomes; therefore, we examined statistical interaction between dosage and duration adequacy separately using both a multiplicative interaction term and a 4-part dummy adherence variable. The 4-part variable was coded as all possible unions of dosage and duration adequacy versus inadequate dosage and inadequate duration (referent group). We did not observe statistical interaction; therefore, the simpler models are presented here (Table 2).

Subjects with missing data were excluded from all regression analyses. Significant differences were demonstrated between the analytic subsample (those with data entered into the multivariable models) and the overall sample. The overall sample was 1 year older ($P < 0.05$), had a $1500 higher annual income ($P < 0.05$), and was 3% more married ($P < 0.05$) compared with the analytic subsample.

We were concerned that hospitalization during the profiling period might have affected duration adequacy (ie, medications could be temporarily discontinued during hospitalization); consequently, we used a chi-squared test to compare duration adequacy between those who were hospitalized during the profiling period (8.3% of the sample) and those who were not hospitalized. We identified a statistically significant difference in depression care adequacy: 41% duration adequacy for those hospitalized during the profiling period versus 45% for those without hospitalizations during the profiling period. Although this is probably not a clinically meaningful difference, we distinguished hospitalization during the prior 12 months from hospitalization occurring in the profiling period in assessing its relationship with subsequent overall and subsequent psychiatric hospitalization (Table 2).

We were concerned that the link between depression care process and outcome might be moderated by depression stage and/or depression diagnosis type. Therefore, we separately examined the association between depression care and subsequent hospitalization for each depression stage and depression diagnosis type (ie, major depression vs. depressive disorder NOS). We did not identify a significant association after stratifying the analyses by depression diagnosis type (data not shown). Additional analyses of the depression diagnosis types have been reported elsewhere.[30] We adjusted for follow-up visit adequacy (Y/N) in the multivariable regressions for the acute stage. We did not identify a significant association; therefore, we excluded follow-up visit adequacy from the final models (Table 3).

All analyses were performed using SAS version 8.0 (SAS Institute Inc., Cary, NC).

RESULTS

We identified 2399 patients (19%) who had at least one hospitalization during the 12 months after the profiling period. Of these, 965 patients had at least 1 psychiatric hospitalization. Overall, 1272 patients (53%) had only 1 hospitalization; the remainder had 2 or greater. The majority of hospitalizations were for acute care (74%), 11% were for observational care, and the remainder for extended care (data not shown). PTSD and alcoholism were the leading causes of hospitalization (20% overall). Three percent of the hospitalizations were the result of depression (10th leading cause).

TABLE 2. Association of Depression Care Adequacy With Subsequent Hospitalization

Independent Variables	Hospitalization in 12 Months After Depression Care Profiling Period (Y/N)*	
	Overall Hospitalization OR (95% CI) (N = 10,453)	Psychiatric Hospitalization OR (95% CI) (N = 10,453)
Age		
≤65 years	.87 (.77–.98);	2.29 (1.77–2.97)
≥65 years	1 (referent)	1 (referent)
Sex		
Female	.88 (.72–1.09);	.67 (.49–.97)
Male	1	1
Race		
Black	1.35 (1.11–1.64);	1.55 (1.20–1.98)
White	1	1
Marital status		
Not Married	1.07 (.97–1.20);	1.21 (1.01–1.43)
Married	1	1
Annual income		
<$40,000	1.04 (.93–1.16);	1.13 (.95–1.34)
≥$40,000	1	1
Hospitalization in prior 12 months		
Profiling period only, Y/N	4.54 (3.70–5.39);	4.63 (3.54–6.06)
Intervening 9 months only, Y/N	3.43 (3.02–3.90)	3.59 (2.96–4.36)
Intervening and profiling, Y/N	11.18 (9.11–13.72)	10.78 (8.44–13.76)
No hospitalization	1	1
Comorbidity index (continuous variable, 0–30 conditions)	1.19 (1.16–1.23);	.89 (.85–.93)
Anxiety disorder		
Y vs. N	1.07 (.97–1.19);	1.11 (.95–1.31)
Alcoholism		
Y vs. N	1.47 (1.30–1.67);	2.89 (2.41–3.45)
PTSD		
Y vs. N	1.11 (.99–1.24);	1.67 (1.41–1.96)
Dosage adequacy		
Y vs. N	1.12 (.94–1.34)	1.09 (.84–1.44)
Duration adequacy		
Y vs. N	.90 (.81–1.00)	.82 (.69–.96)

*These data reflect 2 separate multivariable regression models: 1 for subsequent overall hospitalization and 1 for subsequent psychiatric hospitalization. OR = odds ratio; CI = confidence interval; PTSD = posttraumatic stress disorder.

We presented depression care adequacy results in full in a prior paper.[30] In short, we identified 90% dosage adequacy and 45% duration adequacy in the overall sample. We identified 62% follow-up visit adequacy in the acute stage group (Table 1). Of the overall sample, we characterized 6% (778 patients) as acute stage, 17% (2095 patients) as continuation stage, and 77% (9805 patients) as maintenance stage.

Duration adequacy was identified as a significant predictor of both subsequent overall and psychiatric hospitalization (Table 2). Dosage adequacy was not a significant predictor of either subsequent overall or psychiatric hospitalization. Those who were younger than 65 years, of black race, not married, hospitalized in the prior 12 months, and had comorbid alcoholism and PTSD were more likely to have a subsequent psychiatric hospitalization (Table 2). Those who were of black race, hospitalized in the prior 12 months, had a higher number of comorbid medical illnesses, and comorbid alcoholism and PTSD were more likely to have a subsequent overall hospitalization (Table 2).

TABLE 3. Association of Stage-Specific Depression Care Adequacy With Subsequent Psychiatric Hospitalization

Independent Variables	Psychiatric Hospitalization in 12 Months After Depression Care Profiling Period (Y/N)*		
	Acute Stage OR (95% CI) (N = 683)	Continuation Stage OR (95% CI) (N = 1637)	Maintenance Stage OR (95% CI) (N = 8133)
Age			
<65 years	2.29 (.85–6.13);	3.75 (1.60–8.79);	1.92 (1.47–2.48)
≥65 years	1 (referent)	1 (referent)	1 (referent)
Sex			
Female	.22 (.03–1.76);	.93 (.39–2.22);	.81 (.57–1.16)
Male	1	1	1
Race			
Black	2.32 (.98–5.45);	1.13 (.58–2.21);	1.57 (1.18–2.08)
White	1	1	1
Marital status			
Not married	.94 (.48–1.85);	1.37 (.86–2.19);	1.25 (1.02–1.51)
Married	1	1	1
Annual income			
<$40,000	.99 (.51–1.92);	1.78 (1.06–2.94);	1.02 (.84–1.22)
≥$40,000	1	1	1
Hospitalization in prior 12 months			
Profiling period only, Y/N	5.71 (2.50–13.01);	3.44 (1.19–9.96);	5.46 (4.09–7.28)
Intervening 9 months only, Y/N	2.74 (1.15–6.53);	3.34 (2.04–5.45)	3.92 (3.17–4.85)
Intervening and profiling, Y/N	14.83 (6.21–35.47);	19.25 (10.16–36.44)	12.96 (9.86–17.03)
No hospitalization	1	1	1
Comorbidity index (continuous variable, 0–30 conditions)	.86 (.71–1.04);	.89 (.78–1.02);	.91 (.86–.95)
Anxiety disorder			
Y vs. N	1.41 (.76–2.61);	.99 (.65–1.52);	1.23 (1.03–1.46)
Alcoholism			
Y vs. N	2.17 (1.08–4.34);	3.65 (2.24–5.92);	2.91 (2.40–3.53)
PTSD			
Y vs. N	1.56 (.85–2.87);	2.30 (1.47–3.59);	1.53 (1.28–1.83)
Dosage adequacy			
Y vs. N	1.53 (.66–3.53)	1.39 (.59–3.29)	1.02 (.76–1.37)
Duration adequacy			
Y vs. N	1.21 (.64–2.26)	.86 (.56–1.33)	.83 (.70–.99)

*These data reflect three different multivariable models: one for each of the depression stages (ie, acute, continuation, and maintenance).
OR = odds ratio; CI = confidence interval; PTSD = posttraumatic stress disorder.

In Table 3, we present the results of the depression stage-specific multivariable regression models evaluating the association of depression care adequacy with subsequent psychiatric hospitalization. We do not present the depression stage-specific regression models for subsequent overall hospitalization. The results for these models were similar to those of the overall model presented in Table 2. Of these models, the only significant depression care process–outcome link was within the maintenance stage; those with adequate duration were 17% less likely to have an ensuing hospitalization (OR, .83; 95% CI, .70–.99) than those with inadequate duration.

DISCUSSION

In this study, we used VHA administrative and centralized pharmacy records to identify a population at risk for inadequate care, to define guideline-based depression process measures, and to measure the quality of depression care. This

is of considerable importance to the VHA, which is the largest integrated mental healthcare system in the United States with generous access for their patient population, known to suffer with a higher-than-average prevalence of mental illness.[39] We seek to link depression care process with an important general healthcare outcome, subsequent overall and psychiatric hospitalization. Linking healthcare process and outcomes is critical in demonstrating the validity of quality indicators and the magnitude of benefit from high-quality care.[21] It also augments credibility for practitioners and systems implementing quality improvement programs.

We identified a significant association between 1 depression process measure (duration adequacy) and hospitalization during the 12 months after depression care profiling; those with adequate duration of antidepressants had a 10% lower odds of overall hospitalization and an 18% lower odds of psychiatric hospitalization than those with inadequate duration of antidepressants. This is consistent with the findings of others,[26,28] who have demonstrated that untreated and undertreated depression is a risk factor for higher health services utilization. A similar evaluation has been done recently among a schizophrenic cohort, which illustrated that inadequate care as measured by guideline-based process measures resulted in increased subsequent hospitalization.[34]

We did not demonstrate a significant association between dosage adequacy and subsequent hospitalization; this could be the result of the lack of variation in this measurement (the overwhelming majority had adequate dosing of antidepressants) or the possibility that some were prescribed lower antidepressant doses for therapeutic indications other than depression. We have performed a number of sensitivity analyses to further explore this possibility; these are reported elsewhere,[30] in which we concluded that despite the limited clinical detail of our data sources, they allowed for valuable insight into the process of depression care. Nonetheless, observing a significantly lower odds of subsequent overall and psychiatric hospitalization among those with adequate duration of antidepressants strengthens the predictive validity of this process measure. Further evaluation of the association between appropriate depression care and ensuing healthcare utilization is needed.

We identified suboptimal follow-up visits (62%) among the acute stage. Although we did not demonstrate an association of follow-up visit adequacy with subsequent overall or psychiatric hospitalization, further exploration among a larger sample is warranted. Additionally, future research might explore and define optimal follow-up visit patterns that link to important healthcare outcomes for continuation and maintenance-stage patients.

We present our analyses of the depression care process–outcome link stratified by depression stage in Table 3. We did not demonstrate marked differences among the groups after adjusting for potential confounders. Nevertheless, we add to a growing literature assessing depression stage differences.[40,41] Although this topic needs further study, health systems implementing depression quality improvement programs might want to consider differentiating among the depression stages.

We identified black race as a predictor of subsequent overall and psychiatric hospitalization. We had previously identified black race as a predictor of inadequate depression care.[30] We recognize that there could be several reasons for these racial differences in depression care process and outcomes that are not captured by these data, including patient preferences, health beliefs, concurrent private sector care, and a small, nonrepresentative black subgroup. Nonetheless, this disparity is consistent with earlier studies, which describe differential receipt of needed procedures and appropriate medication regimens between black and white patients.[42–44] Additionally, racial differences in depression care quality have been identified in 2 recent observational studies.[42,44] Further investigation of racial differences in depression care quality is needed.

Prior hospitalization was identified as a strong predictor of subsequent overall and psychiatric hospitalization. More than 1 hospitalization during the prior 12 months (ie, at least once during the profiling period and intervening 9 months) was the strongest predictor of subsequent overall and psychiatric hospitalization. This is not an unexpected finding that is consistent with a large literature predicting rehospitalization.[33,34]

Those younger than age 65 were 13% less likely to have a subsequent overall hospitalization but were 2.29 times more likely to have a subsequent psychiatric hospitalization. This finding is consistent with prior studies demonstrating a disproportionately higher level of psychiatric illness and poorer mental health-related quality of life among younger compared with older patients in the VHA.[37,39] Those with comorbid alcoholism were more likely to have a subsequent overall and psychiatric hospitalization. Those with comorbid PTSD were more likely to have a subsequent psychiatric hospitalization. Although these findings are not surprising, future studies might further define the impact of depression and comorbid psychiatric and medical illness on healthcare outcomes.

More clinically detailed information would allow us to assess the comprehensiveness of our process measures in capturing the quality of depression care. Change in depressive symptomatology over the profiling period would establish criterion validity of these measures, but our data sources do not provide such information. Furthermore, the magnitude of the association between depression care process and outcomes observed here was small. We recognize the possibility that these administrative data have incorrectly identified patients as depressed. The observed association might have been stronger if we could have eliminated falsely identified patients from the sample. Recent work assessing the link

between administrative data-derived depression diagnoses and criterion depression diagnoses has moved the literature toward consensus on identification of depression using administrative data.[31,45,46] Nonetheless, our exclusive use of administrative and centralized pharmacy records mirrors the practical reality of quality profiling in a naturalistic clinical setting. Health systems should consider systematic, centralized collection of depression-specific data. Further examination of the relationship among depression process measures, general healthcare outcomes, and depression-specific outcomes is necessary.

This study has limitations. Restricting depression care profiling to the Northeastern United States limits generalizability to the national veteran patient population. These sites were chosen for ease of data collection; the VHA national pharmacy database was organized just prior to the study design phase, and access to their data was unclear initially. Nevertheless, these hospitals represent diverse geographic settings and reflect the experience of the national veteran patient population.[35] Furthermore, one should maintain caution in applying these findings to non-VHA healthcare settings, especially given that our predominantly male sample might not generalize to the national population undergoing depression care, the majority of whom are women.[2] Nevertheless, the VHA experience in healthcare quality monitoring can be a leading teaching source for other large, integrated health systems.

A great strength of this study is its intensive use of the rich, multifold data sources of the VHA, which is the largest, integrated healthcare system in the United States. These depression process measures have face validity, because guideline-concordant depression care has been connected with improved outcomes in diverse study settings.[23,25,47] The use of clinical guideline benchmarks in defining depression process measures is demonstrated in this study, although more work is needed for validity and reliability to be firmly established. We recognize depression care that falls short of guideline recommendations, especially regarding duration of therapy. Further investigation is needed to determine reasons for these depression care deficits.

Performance measurement has become a prominent reality in today's changing healthcare system. Accurately and reliably capturing healthcare quality is an important mission. The mental healthcare field has contributed substantially to this critical objective. Notably, recent schizophrenia treatment guidelines[48] have rendered findings from considerable efficacy research into salient treatment recommendations and user-friendly quality measures. The Health Plan Employer Data and Information System (HEDIS), a set of standardized performance measures to which health maintenance organizations are held accountable in accreditation, has become an important information source for mental healthcare quality.[49,50] In fact, our depression process measures differ from those of HEDIS only by our use of a measure for adequacy of antidepressant dosing and by our definition of antidepressant duration adequacy. For example, we dichotomized duration adequacy at 79% compared with HEDIS at 93%. Additionally, the operationalism of these measures could vary among health systems with differing centralized pharmacy data access. The centralized pharmacy data of the VHA is amenable to medication adherence formulations used by others and adapted for depression care in this study.[30]

This study adds to our skill in profiling mental healthcare quality and will hopefully inform healthcare systems moving toward translating efficacious quality improvement interventions into practice.[10,13,18] To strive for powerful methods of surveying and protecting the quality of mental health care is crucial in this era of shifting resources and healthcare reorganization.

ACKNOWLEDGMENTS

The authors thank Jeff Whittle, MD, MPH, Director, Division of General and Geriatric Medicine, University of Kansas Medical Center, for his thorough examination of the revised manuscript.

REFERENCES

1. 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.
2. Goldman LS, Nielsen NH, Champion HC. Awareness, diagnosis, and treatment of depression. *J Gen Intern Med*. 1999;14:569–580.
3. Broadhead WE, Blazer DG, George LK, et al. Depression, disability days, and days lost from work in a prospective epidemiologic survey. *JAMA*. 1990;264:2524–2528.
4. Judd LL, Akiskal HS, Maser JD, et al. A prospective 12-year study of subsyndromal and syndromal depressive symptoms in unipolar major depressive disorders. *Arch Gen Psychiatry*. 1998;55:694–700.
5. Wells KB, Schoenbaum M, Unutzer J, et al. Quality of care for primary care patients with depression in managed care. *Arch Fam Med*. 1999;8:529–536.
6. American Psychiatric Association. Practice guideline for major depressive disorder in adults. *Am J Psychiatry*. 1993;(suppl):S1-26.
7. *Clinical Practice Guidelines for Major Depressive Disorder*. Department of Veterans Affairs; 1997.
8. Depression Guidelines Panel. *Depression in Primary Care*, Vol. 2: Treatment of Major Depression. Pub. No. 93–0551. Rockville, MD: AHCPR; 1993.
9. Depression Guidelines Panel. *Depression in Primary Care*, Vol. 1: Detection and Diagnosis. Pub. No. 93–0550. Rockville, MD: AHCPR; 1993.
10. Wells KB, Sherbourne C, Schoenbaum M, et al. Impact of disseminating quality improvement programs for depression in managed primary care: a randomized controlled trial. *JAMA*. 2000;283:212–220.
11. Rost KM, Duan N, Rubenstein LV, et al. The Quality Improvement for Depression collaboration: general analytic strategies for a coordinated study of quality improvement in depression care. *Gen Hosp Psychiatry*. 2001;23:239–253.
12. Meredith LS, Jackson-Triche M, Duan N, et al. Quality improvement for depression enhances long-term treatment knowledge for primary care clinicians. *J Gen Intern Med*. 2000;15:868–877.
13. 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.

14. US Preventive Services Task Force. Screening for depression: recommendations and rationale. *Ann Intern Med.* 2002;136:760–776.
15. Pignone MP, Gaynes BN, Rushton JL, et al. Screening for depression in adults: a summary of the evidence for the US Preventive Services Task Force. *Ann Intern Med.* 2002;136:765–776.
16. Chassin MR, Galvin RW. The urgent need to improve health care quality. Institute of Medicine National Roundtable on Health Care Quality. *JAMA.* 1998;280:1000–1005.
17. Von Korff M, Katon W, Unutzer J, et al. Improving depression care: barriers, solutions, and research needs. *J Fam Pract.* 2001;50:E1.
18. 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.
19. Rost K, Nutting P, Smith J, et al. Improving depression outcomes in community primary care practice: a randomized trial of the quEST intervention. Quality Enhancement by Strategic Teaming. *J Gen Intern Med.* 2001;16:143–149.
20. Solberg LI, Korsen N, Oxman TE, et al. The need for a system in the care of depression. *J Fam Pract.* 1999;48:973–979.
21. Donabedian A. Evaluating the quality of medical care. *Milbank Q.* 1966;44(suppl):166–206.
22. Donabedian A. The quality of care. How can it be assessed? *JAMA.* 1988;260:1743–1748.
23. Melfi CA, Chawla AJ, Croghan TW, et al. The effects of adherence to antidepressant treatment guidelines on relapse and recurrence of depression. *Arch Gen Psychiatry.* 1998;55:1128–1132.
24. Goethe JW, Szarek BL, Cook WL. A comparison of adequately vs. inadequately treated depressed patients. *J Nerv Ment Dis.* 1988;176: 465–470.
25. Fortney J, Rost K, Zhang M, et al. The relationship between quality and outcomes in routine depression care. *Psychiatr Serv.* 2001;52:56–62.
26. Pearson SD, Katzelnick DJ, Simon GE, et al. Depression among high utilizers of medical care. *J Gen Intern Med.* 1999;14:461–468.
27. Croghan TW, Kniesner TJ, Melfi CA, et al. Effect of antidepressant choice on the incidence and economic intensity of hospitalization among depressed individuals. *Admin Policy Ment Health.* 2000;27:183–195.
28. Bula CJ, Wietlisbach V, Burnand B, et al. Depressive symptoms as a predictor of 6-month outcomes and services utilization in elderly medical inpatients. *Arch Intern Med.* 2001;161:2609–2615.
29. Rost K, Smith GR, Burnam MA, et al. Measuring the outcomes of care for mental health problems. The case of depressive disorders. *Med Care.* 1992;30(suppl):MS266-273.
30. Charbonneau A, Rosen AK, Ash AS, et al. Measuring the quality of depression care in a large integrated health system. *Med Care.* 2003;41: 669–680.
31. Kerr EA, McGlynn EA, Van Vorst KA, et al. Measuring antidepressant prescribing practice in a health care system using administrative data: implications for quality measurement and improvement. *Jt Comm J Qual Improv.* 2000;26:203–216.
32. Melfi CA, Croghan TW. Use of claims data for research on treatment and outcomes of depression care. *Med Care.* 1999;37(suppl):AS77-80.
33. Coleman EA, Wagner EH, Grothaus LC, et al. Predicting hospitalization and functional decline in older health plan enrollees: are administrative data as accurate as self-report? *J Am Geriatr Soc.* 1998;46:419–425.
34. Valenstein M, Copeland LA, Blow FC, et al. Pharmacy data identify poorly adherent patients with schizophrenia at increased risk for admission. *Med Care.* 2002;40:630–639.
35. Perlin J, Kazis L, Skinner K, et al. *Health Status and Outcomes of Veterans: Physical and Mental Component Summary Scores Veterans SF-36. 1999 Large Health Survey of Veteran Enrollees. Executive Report.* Washington, DC; Bedford, MA: Department of Veterans Affairs, Veterans Health Administration, Office of Quality and Performance and VHA Health Assessment Project, Center for Health Quality, Outcomes and Economic Research; May 2000.
36. Tarlov AR, Ware JEJr, Greenfield S, et al. The Medical Outcomes Study. An application of methods for monitoring the results of medical care. *JAMA.* 1989;262:925–930.
37. Kazis LE, Miller DR, Clark J, et al. Health-related quality of life in patients served by the Department of Veterans Affairs: results from the Veterans Health Study. *Arch Intern Med.* 1998;158:626–632.
38. Melfi CA, Chawla AJ, Croghan TW, et al. The effects of adherence to antidepressant treatment guidelines on relapse and recurrence of depression. *Arch Gen Psychiatry.* 1998;55:1128–1132.
39. Hankin CS, Spiro A III, Miller DR, et al. Mental disorders and mental health treatment among U. S. Department of Veterans Affairs outpatients: the Veterans Health Study. *Am J Psychiatry.* 1999;156:1924–1930.
40. Reimherr FW, Amsterdam JD, Quitkin FM, et al. Optimal length of continuation therapy in depression: a prospective assessment during long-term fluoxetine treatment. *Am J Psychiatry.* 1998;155:1247–1253.
41. McGrath PJ, Stewart JW, Petkova E, et al. Predictors of relapse during fluoxetine continuation or maintenance treatment of major depression. *J Clin Psychiatry.* 2000;61:518–524.
42. Melfi CA, Croghan TW, Hanna MP, et al. Racial variation in antidepressant treatment in a Medicaid population. *J Clin Psychiatry.* 2000; 61:16–21.
43. Petersen LA, Wright SM, Peterson ED, et al. Impact of race on cardiac care and outcomes in veterans with acute myocardial infarction. *Med Care.* 2002;40(suppl):I86–96.
44. Young AS, Klap R, Sherbourne CD, et al. The quality of care for depressive and anxiety disorders in the United States. *Arch Gen Psychiatry.* 2001;58:55–61.
45. Valenstein M, Ritsema T, Green L, et al. Targeting quality improvement activities for depression. Implications of using administrative data. *J Fam Pract.* 2000;49:721–728.
46. Spettell CM, Wall TC, Allison J, et al. Identifying physician-recognized depression from administrative data: consequences for quality measurement. *Health Serv Res.* 2003;38:1081–1102.
47. Revicki DA, Simon GE, Chan K, et al. Depression, health-related quality of life, and medical cost outcomes of receiving recommended levels of antidepressant treatment. *J Fam Pract.* 1998;47:446–452.
48. Buchanan RW, Kreyenbuhl J, Zito JM, et al. The schizophrenia PORT pharmacological treatment recommendations: conformance and implications for symptoms and functional outcome. *Schizophr Bull.* 2002;28: 63–73.
49. McGlynn EA. Choosing chronic disease measures for HEDIS: conceptual framework and review of seven clinical areas. *Manag Care Q.* 1996;4:54–77.
50. Druss BG, Miller CL, Rosenheck RA, et al. Mental health care quality under managed care in the United States: a view from the Health Employer Data and Information Set (HEDIS). *Am J Psychiatry.* 2002; 159:860–862.

Appendix A Antidepressant Guideline-Recommended Minimum Daily Dosages[7]

Antidepressants	Elderly Dosing (≥65 years)	Nonelderly Dosing (<65 years)
Sertraline	25 mg/day	50 mg/day
Paroxetine	10 mg/day	20 mg/day
Citalopram*	20 mg/day	20 mg/day
Fluoxetine	10 mg/day	20 mg/day
Fluvoxamine*	100 mg/day	100 mg/day
Amitriptyline	Not recommended (0 mg/day)	50 mg/day
Desipramine	25 mg/day	100 mg/day
Doxepin	30 mg/day	75 mg/day
Imipramine	30 mg/day	75 mg/day
Nortriptyline	30 mg/day	75 mg/day
Protriptyline	15 mg/day	15 mg/day
Trimipramine	50 mg/day	75 mg/day
Clomipramine	75 mg/day	75 mg/day
Trazodone	25 mg/day	150 mg/day
Bupropion	50 mg/day	200 mg/day
Phenelzine	45 mg/day	45 mg/day
Tranylcypromine	20 mg/day	20 mg/day
Maprotiline	25 mg/day	75 mg/day
Nefazodone	100 mg/day	200 mg/day
Venlafaxine	50 mg/day	75 mg/day
Mirtazapine	15 mg/day	15 mg/day
Amoxapine	50 mg/day	100 mg/day

*These medications were not addressed in the 1997 VHA Depression Guidelines; minimum daily dosage derived from expert consensus.

ELSEVIER

Contraception

Contraception 69 (2004) 43–45
Original research article

Emergency contraception: prescribing practices of general internists compared with other primary care physicians

C.H. Chuang[a,*], L.J. Waldman[b], K.M. Freund[c], A.S. Ash[a]

[a]*Section of General Internal Medicine, Boston University Medical Center, 720 Harrison Avenue, Suite 1108, Boston, MA 02118, USA*
[b]*Department of Medicine, Emory University School of Medicine, Atlanta, GA, USA*
[c]*Women's Health Unit, Section of General Internal Medicine, Boston University Medical Center, Boston, MA, USA*

Received 30 July 2003; received in revised form 10 September 2003; accepted 15 September 2003

Abstract

Primary care physicians of all specialties should be familiar with prescribing emergency contraception (EC). We conducted a mail survey of 282 randomly sampled physicians in general internal medicine (31%), family medicine (34%) and obstetrics-gynecology (35%). Experience with prescribing EC significantly differed by specialty (63% of general internists, 76% of family physicians, and 94% of obstetrician-gynecologists, $p < 0.0001$). Controlling for year of graduation, gender, religion and practice location, family physicians [adjusted odds ratio (OR): 2.5, 95% confidence interval (CI): 1.2–5.2] and obstetrician-gynecologists (adjusted OR: 11.2, 95% CI: 4.0–31.3) were still significantly more likely to have ever prescribed EC than general internists. Efforts to increase awareness and knowledge of EC should be aimed at general internists since they provide primary care for many reproductive age women.

Keywords: Emergency contraception; Practice patterns

1. Introduction

Widespread availability of emergency contraception (EC) could have a tremendous impact in reducing the number of unintended pregnancies and abortions in this country [1]. EC pills, which are combined estrogen-progestin pills or progestin-only pills, are safe and effective for preventing unintended pregnancies. They are more effective when taken as soon as possible after unprotected intercourse. Therefore, it is critical that a woman seeking EC be able to obtain a physician's attention as soon as possible. Obtaining medication can be delayed if the physician is unfamiliar with how to properly evaluate the patient, unable to provide appropriate counseling or uncertain how to prescribe the method.

In order for EC to be accessible, primary care physicians of all specialties need to be familiar with this method. In a 1997 survey, the Kaiser Family Foundation reported that 85% of obstetrician-gynecologists and 50% of family physicians had prescribed EC at least once in the preceding year [2]. Other studies have further characterized the EC prescribing practices of obstetrician-gynecologists, pediatricians, adolescent health specialists and family physicians [2–8]. Although internal medicine physicians provide primary care for a significant proportion of reproductive-age women [9], little is known about their EC prescribing practices. We sought to compare the EC prescribing practices of general internists with those of other primary care providers of reproductive-age women.

2. Methods

2.1. Study subjects

Approval for this study was obtained from the Institutional Review Board for Human Research at the Boston University Medical Center. Board-certified physicians in internal medicine, family medicine and obstetrics-gynecology were identified from the Folio's of Massachusetts Database, a listing of registered physicians in Massachusetts. We retained only physicians who were currently practicing

* Corresponding author. Tel.: +1-617-638-7515; fax: +1-617-638-8026.
E-mail address: Cynthia.chuang@bmc.org (C.H. Chuang).

0010-7824/04/$ – see front matter
doi:10.1016/j.contraception.2003.09.003

Table 1
Physician characteristics

Characteristic	General internists (n = 86)	Family physicians (n = 96)	Obstetrician-gynecologists (n = 100)	p Value[a]
	n (%)	n (%)	n (%)	
Year of medical school graduation				
Before 1980	39 (45)	39 (41)	62 (62)	<0.01
Gender				
Male	57 (66)	69 (72)	61 (61)	0.27
Religion				
Catholic	15 (17)	29 (30)	30 (30)	0.08
Practice location				
Urban	43 (51)	29 (31)	38 (38)	0.03
Suburban	35 (41)	47 (50)	52 (52)	
Rural	7 (8)	18 (19)	10 (10)	
Contraceptive provider for >25% of female patients of reproductive age	45 (52)	67 (70)	89 (89)	<0.0001

[a] χ^2 test for homogeneity.

in Massachusetts, not currently in training and not listed as subspecialists. We randomly sampled 200 such physicians from each of the three specialties, resulting in 600 potential study subjects.

2.2. Survey instrument

The 29-item survey was developed to assess physician attitudes and practice patterns regarding certain reproductive health issues. A question assessing EC practices asked: "As part of your clinical practice, how often do you prescribe emergency contraception? (e.g., Ovral 2 pills every 12 hours for 2 doses)?" The categorical response choices were "never," "1–5 times/year," "6–10 times/year," and ">10 times/year." Respondents also supplied information on demographics (specialty, year of medical school graduation, gender and religion) and practice characteristics (practice location and contraception prescribing practices). The survey was initially piloted to 10 physicians representing the three different specialties. The anonymous questionnaire was then mailed to the 600 physicians in April 2000.

2.3. Analysis

All statistical analyses were performed using SAS for Windows, Version 8.0 (SAS Institute, Cary, NC, USA). The main outcome variable was dichotomous: "ever prescribed emergency contraception" versus "never prescribed emergency contraception." Multivariable analysis was performed using logistic regression to examine the simultaneous effects of physician specialty, year of medical school graduation, gender, religion and practice location.

To further assess prescribing frequencies, EC prescribing rates were then categorized into 0–5 times per year vs. >5 times per year. The percentages of physicians who prescribed at these different frequencies were then compared across specialties. All analyses were two-tailed, using $p < 0.05$ as criterion for statistical significance.

3. Results

Of the 600 physicians who were mailed the survey, 55 replied that they were no longer practicing in Massachusetts and 33 questionnaires were returned because of incorrect addresses, leaving 512 subjects eligible for the study. Three-hundred completed surveys were returned for a response rate of 59%. We then excluded 16 respondents because they were either subspecialists or did not indicate their specialty. Another two physicians were excluded because they did not answer the question on EC. The remaining 282 subjects were analyzed. There were 86 internists (31%), 96 family physicians (34%), and 100 obstetrician-gynecologists (36%) in the final sample (Table 1). The specialties differed with respect to several characteristics. The obstetrician-gynecologists were more likely to have graduated from medical school before 1980 (62%) compared with the internists (45%) and the family physicians (41%, $p = 0.007$). The majority of internists worked in urban neighborhoods (51%), while the family physicians and obstetrician-gynecologists were more likely to have suburban or rural practices (69%, and 62%, respectively, $p = 0.03$). Although most of the internists prescribe contraception for at least 25% of their female patients of reproductive age (52%), the family physicians (70%) and obstetrician-gynecologists (89%) were more likely to do so ($p < 0.0001$).

The specialties differed substantially in their EC prescribing practices: family physicians (76%) and obstetrician-gynecologists (94%) were more likely to report having ever prescribed EC than general internists (64%, $p < 0.0001$). In the multivariable analysis adjusting for year of medical school graduation, gender, religion and practice location (Table 2), family physicians [adjusted odds ratio

Table 2
Analyses for ever prescribing emergency contraception by specialty

Characteristic	Ever prescribed	Crude OR	Adjusted OR[a]
	n (%)	(95% CI)	(95% CI)
Specialty			
General internists	55 (64)	ref[b]	ref[b]
Family physicians	73 (76)	1.8 (0.9–3.4)	2.5 (1.2–5.2)
Obstetrician-gynecologists	94 (94)	8.8 (3.5–22.5)	11.2 (4.0–31.3)

OR = odds ratio; CI = confidence interval.
[a] Adjusted for year of medical school graduation, gender, religion and practice location.
[b] General internists served as reference group.

(OR): 2.5, 95% confidence interval (CI): 1.2–5.2] and obstetrician-gynecologists (adjusted OR: 11.2, 95% CI: 4.0–31.3) remained more likely to report having ever prescribed EC than general internists. Being a female physician was a positive predictor (adjusted OR: 9.6, 95% CI: 3.2–29.1), while Catholic religion was a negative predictor (adjusted OR: 0.39, 95% CI: 0.19–0.79) for having ever prescribed EC.

When examining prescribing frequency, we found that 75% of the respondents (86% of general internists, 82% of family physicians and 57% of obstetricians-gynecologists) report prescribing EC no more than five times a year. Therefore, the majority of the physicians, regardless of specialty, prescribe EC infrequently. Only 10% of all the physicians reported prescribing more than 10 times a year.

4. Discussion

Although general internists provide over 20% of all non-obstetric outpatient care to women of reproductive age [9], their EC prescribing practices have not been described previously. In this anonymous, self-administered survey, general internists were significantly less likely to report ever prescribing EC than family physicians and obstetrician-gynecologists. This finding persisted when controlling for other potential predictors of prescribing EC.

Strengths of this study were the use of a random sampling method and the ability to control for several potential confounders. The study also has several limitations. Although a diverse group of practitioners from around the state with a wide range of ages and practice settings were sampled, Massachusetts' physicians may not represent physicians nationwide. Another concern is respondent bias, where respondents willing to complete a survey about reproductive health issues may have more strongly held beliefs about contraceptive management and be either more or less likely to prescribe EC than nonresponding physicians. As with any self-report study, there may be reporting bias, where a physician may report what they feel is ideal behavior, rather than true behavior.

Another limitation is differences in the characteristics of the specialties studied. Obstetrician-gynecologists are likely to care for more reproductive-age women than the other specialties. Also, women with multiple providers may preferentially request EC from their obstetrician-gynecologists rather than from their internists. Our survey did not inquire about the number of requests that physicians received for EC. Since internists and family physicians may be faced with fewer opportunities to prescribe EC, we chose whether EC had ever been prescribed as the main outcome variable. Therefore, any single experience with providing EC was considered, in an attempt to reduce bias due to differences in frequency of requests.

Primary care providers of all specialties should provide routine contraceptive care, including the provision of EC. Although the family physicians and obstetrician-gynecologists prescribed EC more frequently than general internists, the majority of physicians in all three specialties reported prescribing less than five times a year. Physicians are likely missing many opportunities for counseling and prescribing EC. Providers should to be prepared to urgently provide EC to the patient who presents after unprotected intercourse. They should also provide advance prescriptions during routine visits. This study suggests that general internists prescribe EC less than other primary care physicians. Efforts should be aimed at general internists to increase their awareness of EC.

References

[1] Trussell J, Stewart F. The effectiveness of postcoital hormonal contraception. Fam Plann Perspect 1992;24:262–4.
[2] Delbanco SF, Stewart FH, Koenig JD, Parker ML, Hoff T, McIntosh M. Are we making progress with emergency contraception? Recent findings on American adults and health professionals. J Am Med Womens Assoc 1998;53:242–6.
[3] Gold MA, Schein A, Coupey SM. Emergency contraception: a national survey of adolescent health experts. Fam Plann Perspect 1997;29:15–9, 24.
[4] Golden NH, Seigel WM, Fisher M, et al. Emergency contraception: pediatricians' knowledge, attitudes, and opinions. Pediatrics 2001;107:287–92.
[5] Grossman RA, Grossman BD. How frequently is emergency contraception prescribed? Fam Plann Perspect 1994;26:270–1.
[6] Lindsey JN. Emergency contraception. Attitudes and practices of primary care doctors in North Carolina. N C Med J 2001;62:158–62.
[7] Sills MR, Chamberlain JM, Teach SJ. The associations among pediatricians' knowledge, attitudes, and practices regarding emergency contraception. Pediatrics 2000;105:954–6.
[8] Beckman LJ, Harvey SM, Sherman CA, Petitti DB. Changes in providers' views and practices about emergency contraception with education. Obstet Gynecol 2001;97:942–6.
[9] Bartman B, Weiss K. Women's primary care in the United States: a study of practice variation among physician specialties. J Womens Health 1993;2:261–8.

0145-6008/04/2802-0313$03.00/0
ALCOHOLISM: CLINICAL AND EXPERIMENTAL RESEARCH

Vol. 28, No. 2
February 2004

The Veterans Aging Cohort Study: Observational Studies of Alcohol Use, Abuse, and Outcomes Among Human Immunodeficiency Virus–Infected Veterans

Joseph Conigliaro, Tamra Madenwald, Kendall Bryant, Scott Braithwaite, Adam Gordon, Shawn L. Fultz, Stephen Maisto, Jeffrey Samet, Kevin Kraemer, Robert Cook, Nancy Day, Diedra Roach, Susan Richey, and Amy Justice

This article represents the proceedings of a symposium at the 2003 annual meeting of the Research Society on Alcoholism in Fort Lauderdale, FL. The organizers/chairs were Joseph Conigliaro and Amy Justice. The presentations were (1) Introduction, by Joseph Conigliaro and Tamra Madenwald; (2) Alcohol and HIV/AIDS: the importance of integrative and translational research, by Kendall Bryant; (3) Alcohol use and abuse among patients with HIV infection, by Joseph Conigliaro and Stephan Maisto; (4) Severity of comorbid alcohol use/abuse in HIV infection, by Amy Justice and Jeffrey Samet; (5) Estimating the impact of alcohol use on long-term HIV outcomes, by Scott Braithwaite and Amy Justice; (6) Homelessness, drug & alcohol use among HIV+ veterans, by Adam Gordon and Robert Cook; and (7) Hepatitis C & alcohol in the VACS 3 study, by Shawn Fultz and Kevin Kraemer. The symposium concluded with a discussion led and facilitated by Diedra Roach.

Key Words: HIV/AIDS, Alcohol, Comorbidity, Outcomes, Homelessness.

INTRODUCTION

Joseph Conigliaro and Tamra Madenwald

ALCOHOL USE AND abuse may have significant implications for the clinical management of human immunodeficiency virus (HIV)-infected patients and may play a major role in determining patient outcomes. Alcohol abuse in HIV is believed to be associated with increased morbidity and mortality, more rapid disease progression, poorer adherence to antiretroviral regimens, and greater risk of viral resistance (Conigliaro et al., 2003; Cook et al., 2001; Justice et al., 2002; Kalichman et al., 2000; Wagner et al., 2001). Alcohol's effect on HIV, however, has not yet been fully characterized. To date, HIV cohort studies have neither collected adequate data on alcohol use nor sufficiently considered the effect of the interaction of alcohol use and comorbid disease on patient outcomes. This symposium reviews the research design and findings to date from the Veterans Aging Cohort Studies (VACS) to illustrate the multifaceted role of alcohol in health outcomes among HIV-infected veterans and to highlight the design and potential future analyses as more sophisticated data are collected in the current VACS (Table 1). VACS, a collaboration funded by the NIAAA between the NIAAA and the Department of Veterans Affairs (VA), is a multisite 5-year longitudinal cohort study of the effect of alcohol use and chronic disease on health outcomes in veterans with and without HIV infection. In the long term, these analyses should guide clinical decision-making and inform the development of effective interventions for HIV-infected patients who use and abuse alcohol.

VACS is a longitudinal study of 6000 veterans receiving care at 8 VA medical centers. When completed, the cohort will consist of 3000 HIV-infected and 3000 noninfected veterans who are matched by age, race, and clinical site. The study began in June 2002 and is currently completing baseline enrollment, with data collection continuing through June 2007. As of September 2003, VACS had recruited more than 4700 subjects. VACS was preceded by

From the Center for Health Equity Research and Promotion (CHERP), Pittsburgh, PA (JC, AG, SR); Mental Illness Research Evaluation and Clinical Center (MIRECC), Pittsburgh, PA (JC, AG); VA Pittsburgh Healthcare System, Pittsburgh, Pennsylvania (JC, AG, SR); Center for Research on Health Care (CRHC), Pittsburgh, PA (JC, SB, AG, KK, RC); University of Pittsburgh, Pittsburgh, Pennsylvania (JC, TM, SB, AG, KK, RC, ND, SR); NIAAA, Bethesda, MD (KB, DR); Yale University, New Haven, Connecticut (SLF, AJ); Syracuse University, Syracuse, New York (SM); and Boston University Medical Center, Boston, Massachusetts (JS).

Received for publication September 15, 2003; accepted October 13, 2003.

Supported by a Robert Wood Johnson Foundation Generalist Physician Faculty Scholar Award (036505; AJ) and by an award from the NIAAA (AA 13566-02; AJ). JC received support from a Career Development Award from the HSR&D Service, Department of Veterans Affairs (CD-97324-A), and a Robert Wood Johnson Foundation Generalist Physician Faculty Scholar Award (031500). AJ received support from a K 23 Mentored Career Development Award from the NIA (AG00826-03). AG is a Career Development Awardee from the HSR&D Service, Department of Veterans Affairs (RCD-00038-2).

Reprint requests: Joseph Conigliaro, MD, MPH, VA Pittsburgh Healthcare System, University Drive C, Pittsburgh, PA 15240; Fax: 412-688-6916; E-mail: joseph.conigliaro@med.va.gov.

DOI: 10.1097/01.ALC.0000113414.73220.21

Table 1. Evolution of the VACS Studies

Variable	VACS 3	VACS 5	VACS
n	881	1800	6000 (current *n* = 4700)
Dates	June 1999 to July 2000	October 2001 to June 2002	June 2002 to present
No. sites	3	5	8
HIV-negative controls	No	Yes	Yes
Focus	Outcomes, comorbidity, aging, and provider relationships	Feasibility, neuro psychology, blood testing	Alcohol, other comorbid illness, intervention design

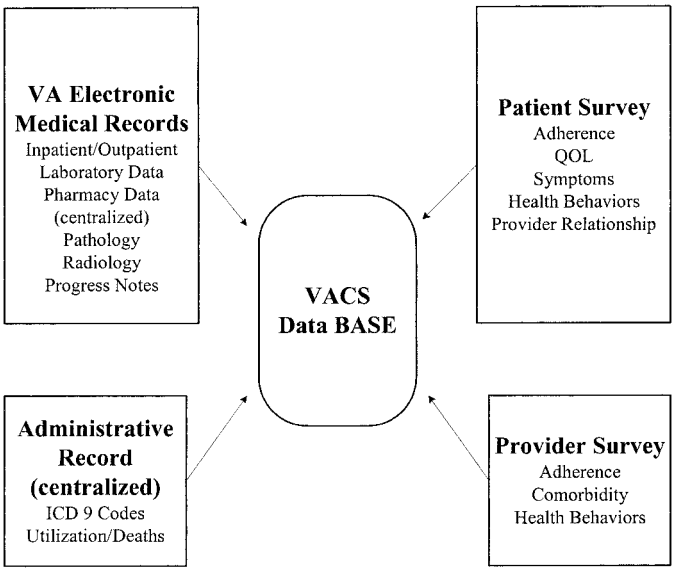


Fig. 1. VACS core data sources.

two smaller studies. VACS 3 (June 1999 though July 2000) collected survey, administrative, and electronic medical record data on 881 HIV-infected veterans receiving care at 3 VA HIV clinics. VACS 5 (October 2001 though June 2002) collected similar feasibility data on 1851 veterans, 1064 HIV-infected veterans, and 787 noninfected veterans who were matched by age, race, and clinical site at 5 VA medical centers.

The collection of patient and provider factors is invaluable and unique to VACS. Although many cohort studies have relied on medical records to collect data, the collection of medical record data is often difficult because much time is spent focusing on finding patients' charts and holding them long enough to extract data. Obtaining relevant laboratory and pharmacy data is also a burden, because often laboratory results are misplaced or not filed at the time the chart is found. As a result, certain critical data elements are missed. However, the VA has the largest electronic medical records system in the world. It is fully integrated and easily accessible nationwide; therefore, VACS is able to collect critical up-to-date data elements from each patient's medical records, such as inpatient and outpatient data; laboratory data; pharmacy data; and pathology, radiology, and progress notes (Fig. 1). By concentrating on linking a larger variety of data sources on individual patients (Fig. 1), VACS is able to measure and test the association of patient factors with HIV measures. These patient factors include medication adherence, quality of life, symptom burden, health behaviors and provider relationships from the patient survey, International Classification of Diseases (ICD)-9 codes and utilization and death records from the VA administrative record, adherence determination, presence of comorbidities, and health behavior and symptoms assessments from the provider survey (Fig. 1).

Obtaining provider data enables VACS researchers to understand the provider's assessment of the severity and prognosis of HIV disease, the presence or absence of comorbid conditions, and the provider's determination of whether the patient is adhering to his or her medication. The provider's awareness of alcohol consumption and related problems, a critical element in VACS, is also determined.

Blood and tissue banking allows VACS researchers to address important underlying mechanistic questions as they are defined by clinical observation and research. These include evaluating surrogate markers, systematically studying pathophysiologic markers of drug and alcohol toxicity, and evaluating genetic predisposition to immune decline.

The specific aim of all the VACS studies (VACS 3, VACS 5, and the current VACS) is to test the association of patient factors, including comorbidities and behaviors, with HIV measures, such as adherence to medications, CD4 cell count, HIV viral load, and drug toxicity, as well as their association with HIV disease progression, symptom burden, quality of life, comorbid disease, and survival. VACS is designed to specifically test the association of alcohol use and abuse with HIV measures. The studies also assess patient and provider attitudes concerning alcohol use and its interaction with HIV infection.

The number and scope of the alcohol measures have differed among the three VACS studies (Table 2). Both VACS 3 and VACS 5 used only the Alcohol Use Disorders Identification Test (AUDIT) (Babor et al., 1992; Conigrave et al., 1995; Fiellin et al., 2000; Saunders and Conigrave, 1990), provider assessments, and ICD-9 codes. VACS, however, makes use of multiple alcohol measures, including the Lifetime Drinking History (Skinner and Sheu, 1982), alcohol timeline follow-back (Sobell and Sobell, 1994), AUDIT, Short Inventory of (alcohol) Problems (Miller et al., 1995), Alcohol Dependence Scale (Skinner and Horn, 1984), and Composite International Diagnostic Interview Substance Abuse Module (Cottler et al., 1989). VACS also includes several measures to evaluate adher-

Table 2. Alcohol-Related Measures in VACS 3, VACS 5, and VACS

Variable	VACS 3	VACS 5	VACS
AUDIT	X	X	X
Provider assessments	X	X	X
ICD-9 codes	X	X	X
Short Inventory of Problems			X
Alcohol Dependence Scale			X
Lifetime Drinking History			X
Timeline follow-back			X
CIDI Substance Abuse module			X
Biomarkers	AST, ALT, MCV	AST, ALT, MCV	AST, ALT, MCV, CDT, C, vitamin B12, folate, mitochondrial toxicity
Alcohol-related focus groups			X

MCV, mean corpuscular volume; CDT, carbohydrate-deficient transferrin; CIDI, Composite International Diagnostic Interview.

ence to medications, including an adapted 30-day timeline follow-back. One aim is to improve existing measures of alcohol use and adherence among HIV-positive patients by determining which assessments are most relevant to HIV-infected patients and patients with other chronic diseases. Future VACS data will continue to include more sophisticated measures of alcohol-use patterns and alcohol diagnoses needed to develop provider- and system-based interventions within this population.

The presentation summaries that follow analyze preliminary data from VACS 3 that will guide the larger study's development as it evolves and that will influence the design of future interventions, particularly as they relate to alcohol use. Kendall Bryant placed the VACS study in the context of translational research and showed how it offers an opportunity to move significant research findings in alcohol and HIV/acquired immune deficiency syndrome (AIDS) effectively into clinical practice. Joseph Conigliaro and Stephan Maisto described levels of excessive alcohol use among HIV-infected veterans and its associated health risks, including patient and provider knowledge of those risks. Amy Justice and Jeffrey Samet examined the association of current hazardous alcohol use and/or past abuse with comorbid conditions and symptom burden. Scott Braithwaite and Amy Justice used a computer simulation to estimate the effect on long-term HIV outcomes of medication nonadherence associated with alcohol use. Adam Gordon and Robert Cook described the prevalence of homelessness, rates of alcohol abuse, and number of physician visits among veterans with HIV. Finally, Shawn Fultz and Kevin Kraemer examined the prevalence of past and current alcohol use among veterans co-infected with hepatitis C virus (HCV) and assessed provider awareness of such use.

ALCOHOL AND HIV/AIDS: THE IMPORTANCE OF INTEGRATIVE AND TRANSLATIONAL RESEARCH

Kendall Bryant

The VACS cohort of veterans represents an important population in which to study alcohol and HIV infection. Among veterans, both alcohol abuse (or dependence) and HIV infection are prevalent; veterans are older and more likely to have co-occurring conditions related to both alcohol and HIV. Among persons infected with HIV, veterans are also more likely to be African American or Hispanic, to have a diverse exposure history, and, as such, to reflect the evolving HIV epidemic.

The overall goal of the VACS is to examine in greater detail the role of alcohol use, abuse, and dependence in the treatment of HIV disease. The relationship between alcohol use, HIV treatment, and progression of disease is complex. Patterns of alcohol use directly and indirectly affect the survival of HIV-infected individuals who drank heavily in the past and may currently drink. Treatment providers often underestimate the cumulative effect of alcohol misuse on long-term health outcomes.

The VACS study examines the prevalence and nature of alcohol consumption and characterizes its associated risk. It also explores patient and provider beliefs about alcohol consumption and behavior change in both cross-sectional and longitudinal research frameworks. The sophisticated VA electronic record systems used in describing patient-provider encounters and additional information on individuals' history of alcohol consumption allows for detailed characterization of alcohol's complex association with HIV disease severity and progression and with other co-occurring disease over time. Finally, the long-term goal of VACS is to test dynamic interventions that involve provider-patient relationships guided by detailed knowledge of these interchanges, which improve outcomes for patients aging with HIV infection.

Although we have described the immediate and future goals for the set of VACS studies (VACS 3, VACS 5, and VACS), the current study (VACS), funded by NIAAA, is intended to provide a platform for developing an integrative approach to research that allows for translation of alcohol and AIDS research into clinical practice. The current study describes the complex relationships of alcohol and HIV and allows for a more complete picture of the patient and clinical setting in which the treatment takes place, which is elaborated from a variety of behavioral and biological scientific perspectives. Thus, the VACS provides a rich source of data and a platform for integrative and cross-disciplinary research in both the behavioral and biological sciences. As a result, for example, the effects of

drinking patterns on adherence to antiretroviral medication and subsequent disease progression (viral replication and immunological sufficiency) can be studied in tandem and integrated with an understanding of liver dysfunction.

In addition, VACS is designed to become an important tool in translational research, a research-to-practice initiative supported by the NIH and the VA that is designed to bridge the gap between researchers and clinicians by (1) identifying research advances ready for widespread adoption, (2) preparing materials that facilitate training and adoption of new techniques, and (3) disseminating knowledge about improved treatment and prevention practices. The effectiveness of the translational research initiative depends on "widening the research footprint" to include more diverse groups of patients and more diverse treatment settings. VACS makes this possible. A panel of alcohol experts oversees the current NIAAA-funded study, which integrates state-of-the art alcohol and HIV measurement. Development and testing of critical alcohol and HIV questions allows for testing and integration of these items into patient interviews, which can guide successful treatment in the future. These interviews and critical reminders can be implemented in the electronic patient management system, which the VA supports, to give provider feedback on patient health status and functions as a structural systems intervention. Finally, these new technologies can be carried into new treatment settings both in the US and abroad.

ALCOHOL USE AND ABUSE AMONG PATIENTS WITH HIV INFECTION

Joseph Conigliaro and Stephan Maisto

We know that general medical health-care providers often miss the opportunity to detect alcohol problems among their patients (Conigliaro et al., 1998). It is not known how well HIV providers address alcohol problems among their patients. In light of the mounting evidence that excessive alcohol use may be important in the management of HIV, the objectives of this analysis were to (1) describe levels of excessive alcohol use (including hazardous use, addiction, and dependence) among a group of HIV-infected veterans cared for at three VA medical centers, (2) determine the association between patterns of alcohol use and laboratory markers of HIV disease progression and antiretroviral drug toxicity, and (3) assess provider awareness of both current and past excessive alcohol use. Finally, we sought to determine the patient characteristics that were associated with provider failure to identify excessive alcohol consumption.

In VACS 3, we assessed alcohol use in several ways. We asked the provider whether the patient "drinks too much alcohol" (mutually exclusive response options were "never," "past," or "present"). This question was structured to reflect the provider's personal definition of "drinking too much" that would reflect his or her opinion, clinical experience, and level of training. We also asked the patient a similar question with the same mutually exclusive response options ("never," "past," or "present"). In addition, we assessed at-risk alcohol use on the basis of the AUDIT (a cutoff score of ≥ 8 out of 40 points detected "at risk alcohol use") (Babor et al., 1992; Isaacson et al., 1994). We examined binge drinking alone by using the third question in the AUDIT, which asks whether the patient has had six or more drinks on one occasion (Bush et al., 1998; Gordon et al., 2001). We extracted ICD-9 codes for diagnoses of alcohol abuse and dependence from the VA electronic medical record by identifying all ICD-9 codes for both inpatient and outpatient care for up to 3 years before the patient baseline survey (from fiscal years 1998–2000 and before July 29, 2000). Laboratory variables [HIV-1 viral load, CD4 cell count, hematocrit, mean corpuscular volume, alanine transaminase (ALT), and aspartate transaminase (AST)] were collected directly from the laboratory data in the electronic medical record for the date closest to the baseline assessment. ALT and AST were standardized to the top normal value at each site.

Using data from VACS 3 (Conigliaro et al., 2003), we found that of 881 HIV-infected veterans (median age, 49 years; 99% male; 54% African American), 20% met criteria for current hazardous drinking by scoring 8 or greater on the AUDIT (Babor et al., 1992; Isaacson et al., 1994). Thirty-three percent were binge drinkers (positive response to the third item on the AUDIT), 32% had a chart ICD-9 alcohol diagnosis over the last 5 years, and 12.5 and 66.7%, respectively, were described by their providers as currently or ever drinking "too much." Hazardous/binge drinkers more often had detectable viral loads ($p < 0.001$). Patients with ICD-9 alcohol diagnoses of alcohol abuse or alcohol dependence more often had increased ALT or AST ($p < 0.02$), anemia ($p < 0.001$), and increased mean corpuscular volume ($p < 0.001$). Providers missed hazardous drinking among those with undetectable viral loads ($p = 0.01$), without HCV ($p = 0.09$), and with normal ASTs ($p = 0.07$) and missed ICD-9 alcohol diagnoses among patients without hepatitis and with CD4 less than 200 cells per microliter.

Overall, we found that among HIV-positive veterans, alcohol use and hazardous drinking are common and are associated with measures of HIV disease progression and/or hepatic comorbidity and anemia. We also found that providers more often missed alcohol problems among patients both with less severe HIV and without evidence of liver disease.

Because alcohol use and abuse are common among HIV-infected individuals and because providers do not routinely identify alcohol use and abuse, providers should routinely screen and counsel HIV-infected patients regarding alcohol problems as part of a standard of care, in an effort to minimize disease progression and bone marrow and hepatic toxicity.

Although interesting and important, the baseline VACS 3 results are limited by a lack of details on alcohol con-

sumption patterns, lifetime alcohol consumption, and alcohol consequences. In the current VACS study, more extensive measures, as described previously, are used. As a result, VACS will focus on more specific categories of alcohol use and abuse in the context of the larger pattern of comorbidity and the interaction of alcohol, HIV infection, lifestyle, and behavior change.

SEVERITY OF COMORBID ALCOHOL USE/ABUSE IN HIV INFECTION

Amy Justice and Jeffrey Samet

Many HIV-infected persons are receiving long-term treatment and, consequently, living long enough to age with the infection. Comorbid disease, defined as any medical or psychiatric condition not known to be caused by HIV infection, and its relationship to past and current alcohol use and abuse have yet to be characterized in HIV infection. Similarly, symptom burden (often a marker of comorbid illness) and its association with alcohol use and abuse has yet to be characterized.

We hypothesized that a diagnosis of past alcohol addiction or abuse and current hazardous drinking would have an additive effect for comorbid disease diagnoses and symptom burden, as follows. First, those free of both a diagnosis of addiction or abuse and current hazardous drinking would have the fewest symptoms and comorbid conditions; second, those with only one of these conditions would have intermediate symptoms and comorbid conditions, and; third, those with both conditions would have the most symptoms and comorbid conditions.

Data on AIDS-defining and comorbid conditions were collected in VACS 3 by using patient and provider report, diagnostic codes, laboratory data, and pharmacy data. We identified the best source for each condition by using a uniform approach. Data on symptoms were collected directly from patients via a self-completed standardized and validated instrument, the HIV Symptom Index (Justice et al., 2001).

Current hazardous drinking (defined as an AUDIT score ≥ 8) was present in 36% of the patients surveyed (23% reported only hazardous drinking, and 13% were found to also have a diagnosis of abuse or dependence). An ICD-9 diagnosis of alcohol abuse or dependence alone was present in 14% of the patients' electronic medical records. Of note, 50% of the patients neither reported current hazardous drinking nor had a diagnosis of abuse or dependence. When we used these to test for an increasing trend of comorbid disease and symptom burden, we found the following. Among physical symptoms (fevers/chills, nausea/vomiting, diarrhea, headache, and muscle pain), there was a statistically significant ($p < 0.03$ for each of the symptoms listed) increasing trend such that those with neither a diagnosis of alcohol abuse or dependence nor current hazardous drinking had the lowest burden of symptoms and such that those with both conditions had the highest burden of symptoms; those with one or the other condition had more intermediate levels of symptom burden. This was not as evident for emotional symptoms (sadness, nervousness, and sleep problems). For these symptoms, patients with neither condition had the lowest frequency of symptoms ($p < 0.03$), but there was not a clear increasing trend among those with one condition versus both conditions. It is interesting to note that diabetes and cancer seemed to decrease in prevalence with additional alcohol-related conditions ($p < 0.02$). HCV and depression increased in prevalence with additional alcohol-related conditions ($p < 0.02$).

We conclude that both a history of alcohol abuse or dependence and hazardous drinking are common among HIV-positive veterans in care and that both of these conditions are associated with increased comorbid disease and symptoms. Although these conditions are not universally additive, there is evidence that their effects may be additive when physical symptoms and liver injury are considered. The role of alcohol in determining outcomes in HIV is likely multifaceted, including the timing and intensity of consumption and the medical and psychological susceptibility of the patient to injury from alcohol. VACS will seek to further understand this relationship by studying patterns of use over time and overlapping and interacting comorbid medical and psychiatric disease.

ESTIMATING THE IMPACT OF ALCOHOL USE ON LONG-TERM HIV OUTCOMES

Scott Braithwaite and Amy Justice

Numerous studies of HIV-infected patients have shown that alcohol use is associated with decreased adherence to antiretroviral therapies (Chesney, 2000; Chesney et al., 2000; Cook et al., 2001; Haubrich et al., 1999a,b; Paterson et al., 2000). Patients who score more than 1 on the CAGE are 3.6 times more likely to miss medication doses (Paterson et al., 2000). Using VACS 3 data, Wagner et al. (2001) found that any alcohol use was related to missed medication doses (odds ratio, 1.9). The effect of alcohol on long-term HIV outcomes, however, is largely unknown. Alcohol-associated nonadherence may have long-term implications because nonadherence is a primary cause of treatment failure. Paterson reported that when adherence was greater than 95%, 22% of patients had a viral rebound compared with when adherence was less than 80%, which was associated with an 80% viral rebound. To model the effect of decreased adherence from alcohol use on clinical outcomes, we developed a second-order Monte Carlo computer simulation that estimates the life expectancy of HIV-infected patients with current therapies. The simulation is probabilistic, and it therefore can mimic some of the heterogeneity seen in clinical populations. All patients were followed up until death, and the risk for death was based on characteristics including age, CD4 count, and viral load. This simulation offers three major advances over other HIV models. First, it represents mutations in the HIV

genome and therefore can predict the time to treatment failure of highly active antiretroviral therapy (HAART) regimens. It represents nonadherence to HAART, which has been shown to be the primary cause of virological failure in numerous patient groups and health-care settings. Finally, it has been calibrated and validated with clinical data, which increases its generalizability. The model explicitly represents the two main phenomena that limit the effectiveness of current therapies: poor adherence and development of phenotypical resistance. Patients progress through sequences of regimens until resistance accrues to all drugs and are at risk for death commensurate with CD4 count and regimen efficacy. The magnitude of the association between alcohol use and nonadherence was based on VACS 3 data (Wagner et al., 2001). *Adherent* was defined as taking all doses in the past 7 days, *nonadherent* as missing all doses in the previous 4 days, and *partially adherent* as neither of these. Times were counted from the start of all treatments.

Our model estimated that adherent patients would have mean life expectancies of 30.7 years, partially adherent patients would have mean life expectancies of 22.0 years, and nonadherent patients would have mean life expectancies of only 6.8 years. In VACS, 471 patients were not using alcohol. Of these patients, 34% were adherent, 59% were partially adherent, and 7% were nonadherent. The 313 alcohol users in VACS had poorer adherence (18, 73, and 9%, respectively). On the basis of these adherence data, our model estimated life expectancies of 25.1 and 26.8 years, respectively, for alcohol users and nonusers with CD4 counts of 500. For patients with CD4 counts of 350, our model estimated that alcohol users and nonusers would have mean life expectancies of 22.2 and 23.9 years, respectively. For patients with CD4 counts of 200, it estimated that alcohol users and nonusers would have mean life expectancies of 18.8 and 20.2 years, respectively.

The computer simulation estimated that the poorer adherence to antiretroviral therapies among alcohol users in VACS 3 would result in more than 1 year of life lost, and this result was robust across a wide range of assumptions regarding clinical stage at the start of therapy. This result considers only the portion of alcohol's effect that is mediated through changes in adherence to antiretroviral therapies (Fig. 2) and does not consider possible direct effects of alcohol on the course of HIV disease or on the risk of death from other causes. Future VACS data will include more detailed information on the quantity and pattern of alcohol use and will enroll more patients. These data will permit future runs of the simulation to consider other important pathways through which alcohol may affect survival in HIV disease.

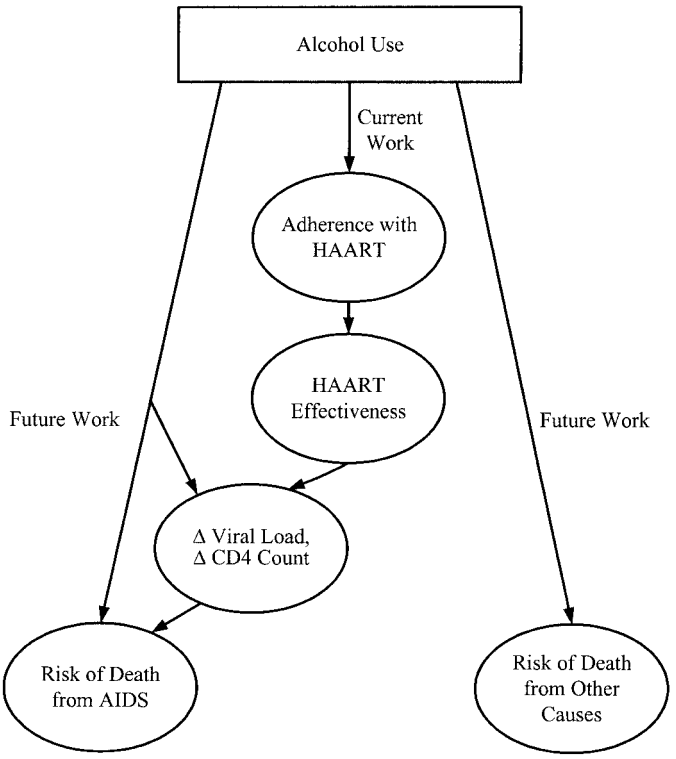


Fig. 2. Relationship of alcohol use to HAART medication adherence and outcomes.

HOMELESSNESS, DRUG & ALCOHOL USE AMONG HIV+ VETERANS

Adam Gordon and Robert Cook

Homelessness and HIV infection are increasingly recognized as important co-occurring conditions. Although estimates vary, as many as 30% of HIV-infected persons may be homeless or marginally housed (Smith et al., 2000). Homelessness may decrease adherence among those with access to HIV care (Wagner et al., 2001), disrupt access to health-care services (Shapiro et al., 1999; Smith et al., 2000), promote sexual and other risky behaviors (Song et al., 2000), and contribute to illicit drug use (Bangsberg et al., 1997; Stein et al., 2000). Alcohol consumption is associated with many of these same outcomes. Homelessness and significant alcohol consumption frequently overlap. However, the relative degree to which homelessness and consumption of alcohol lead to adverse health outcomes in HIV-infected populations is largely unknown. The objectives of this study were to compare, by homeless status, the rates of alcohol and substance use and to explore the association between homelessness, alcohol use, and self-reported utilization of health-care services.

We used data from VACS 3 (described previously), which included two measures of homelessness: "In the past four weeks, have you ever been without a permanent address that you call home?" (defined as current homelessness) and "Have you ever been without a permanent address you call home?" (defined as past homelessness). To determine alcohol use, we defined hazardous drinking as an AUDIT score ≥ 8 (Babor et al., 1992) and binge drinking as any positive answer to the third AUDIT question ("How often do you have 6 or more drinks on one occasion?") (Bush et al., 1998; Gordon et al., 2001). Subjects were also

asked to indicate the number of visits to various health-care providers in the prior 6 months.

For our results, the cohort ($n = 881$) was predominantly male, nonwhite, and middle aged. Overall, 267 (32%) indicated prior homelessness, and of those, 53 (20%) were homeless in the prior 4 weeks. Compared with nonhomeless HIV-infected veterans, homeless HIV-infected veterans were younger, were more likely to be minorities, and were more likely to smoke ($p < 0.05$). Among those who indicated any prior homelessness, 63% drank at hazardous levels by AUDIT criteria, and 35% binge-drank. Compared with persons who were not presently homeless, persons who were homeless in the last 4 weeks ($n = 62$) had similar rates of current alcohol use (35 vs. 41%; $p = 0.4$) and were more likely to be hazardous drinkers by AUDIT criteria (35 vs. 19%; $p = 0.002$) and to be binge drinkers (47 vs. 34%; $p = 0.04$).

Compared with nonhomeless HIV-positive veterans, homeless HIV-positive veterans reported, for the prior 6 months, an increased mean number of visits to any doctor (6.4 vs. 5.5; $p = 0.003$) and to their primary doctor (3.3 vs. 3.0; $p = 0.04$). Among the homeless, similar numbers of visits were reported to any doctor in the prior 6 months between homeless hazardous versus homeless nonhazardous drinkers (6.3 vs. 6.3; $p = 1.0$), and there was a trend toward fewer visits to their primary doctor among binge versus non–binge drinkers (3.6 vs. 4.0; $p = 0.07$).

We conclude that a significant proportion of HIV-infected veterans have a history of homelessness and are currently homeless. Homeless HIV-infected veterans were more likely to drink alcohol at hazardous levels than nonhomeless HIV-infected veterans. Homeless status, but not alcohol use status among the homeless, influenced self-reported utilization of any health-care services and of HIV treatment provider services. Future VACS data will include more detailed information on homelessness, including homeless severity, including typical domiciles (e.g., unsheltered versus emergency sheltered domiciles). The VACS will provide an opportunity to study the interaction of alcohol use disorders and homelessness on a wide range of alcohol-specific, patient-specific, and process outcomes. Until these studies are conducted, HIV providers should be aware of the prominent role of homelessness, and co-occurring alcohol use, in providing care for HIV-infected patients.

HEPATITIS C & ALCOHOL IN THE VACS 3 STUDY

Shawn Fultz and Kevin Kraemer

Hepatitis C (HCV) infection is common in HIV-infected persons (Quan et al., 1993; Sherman et al., 2002; Staples et al., 1999). Because advances in HAART have led to improved survival, chronic HCV infection has the opportunity to progress to clinically significant disease in HIV-infected persons and is emerging as an important cause of morbidity and mortality (Bica et al., 2001; Klein et al., 2003; Monga et al., 2001). For example, chronic HCV liver disease is now the most common cause of non–AIDS-related death among HIV-infected women (Cohen et al., 2002). Despite the considerable burden of disease and the high prevalence of co-infection, factors related to the progression of HCV liver disease in HIV-infected persons are poorly understood. Because alcohol use can accelerate the course of HCV (Poynard et al., 1997), increase the risk of subsequent hepatocellular carcinoma (Khan and Yatsuhashi, 2000), and decrease the efficacy of antiviral treatment (Loguercio et al., 2000), alcohol cessation may be the single most effective intervention among co-infected patients and is strongly recommended (National Institutes of Health Consensus Development Conference Statement, 2002).

Using VACS 3 data, we determined the prevalence of HCV infection and compared the prevalence of current drinking and current hazardous drinking in HIV-infected veterans with and without HCV co-infection. We also determined whether providers were aware of current alcohol and drug use and whether current users were counseled to cut back by their provider. Three hundred (43%) of the 700 HIV-infected veterans tested for HCV were HCV infected. HCV-infected veterans were more likely to be of a minority race and to be injection drug users and were less likely to report having sex with men. Eighty-eight (30%) HCV-infected veterans reported current drinking, compared with 186 (47%) HCV-negative veterans. However, HCV-infected current drinkers were significantly more likely than HCV-negative current drinkers to report an AUDIT score of 8 or higher (51 vs. 26%; $p < 0.05$) and to report binge drinking (66 vs. 52%; $p < 0.05$). Providers recognized only 29 (33%) of the 88 HCV-infected drinkers as currently drinking ($\kappa = 0.07$; sensitivity, 18%; specificity, 93%). Only 25 (28.7%) of the HCV patients who were currently drinking reported that they had been advised to stop, but patients who were HCV positive were more likely to report having been told to cut back on their drinking compared with HCV-negative patients (28.7 and 15.1%, respectively; $p < 0.0005$). Among HCV-infected drinkers, those with an AUDIT score of 8 or greater were significantly more likely to have received advice to cut back than those with scores less than 8 (73 vs. 14%; $p < 0.005$).

These results confirm a high prevalence of HCV co-infection in HIV-infected veterans. Unfortunately, approximately one third of veterans co-infected with HCV and HIV continue to drink alcohol, often at a hazardous level. Continued alcohol use in these co-infected veterans may accelerate the progression of chronic liver disease and adversely affect the efficacy of HAART and interferon-based antiviral HCV therapy. In VACS 3, the problem of continued drinking in veterans co-infected with HCV and HIV was exacerbated by a lack of awareness in health-care providers. As a result, only one quarter of HCV-infected patients who were actively drinking reported being told to stop or cut back. If confirmed in continuing VACS studies,

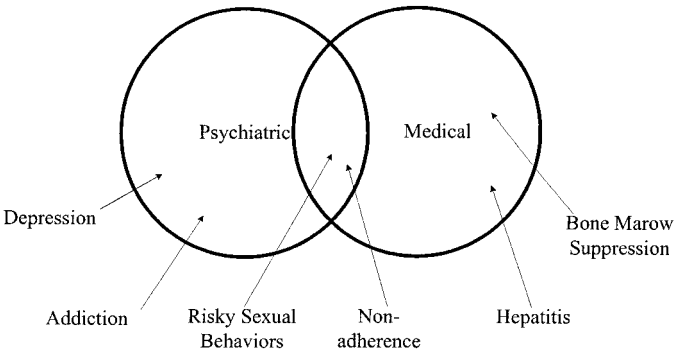


Fig. 3. Alcohol effects on psychiatric and medical outcomes.

this represents an important gap in the provision of quality health care to these patients.

The current VACS study will allow for a more precise quantification of alcohol use and its role in HCV infection among veterans with and without HIV infection. Longitudinal follow-up in VACS will allow us to prospectively measure the effect of alcohol use on hepatotoxicity, other clinical liver disease outcomes, and response and adherence to HAART and antiviral HCV therapy. These data will inform our future efforts to develop provider- and system-based interventions to overcome barriers to alcohol cessation among all HIV-positive patients, especially those co-infected with HCV.

SIGNIFICANCE

Joseph Conigliaro and Tamra Madenwald

With VACS 3 data, alcohol use and hazardous drinking were found to be common in HIV-infected veterans and to be associated with increased symptom burden, more severe comorbid disease, decreased adherence to treatment, and decreased utilization of health services in homeless HIV-infected veterans. In addition, health-care providers were often unaware of alcohol use in HIV-infected veterans, even in the context of patients with a comorbid condition such as HCV infection, for which abstinence is recommended. Similarly, HIV-infected veteran patients who drink reported seldom being counseled to limit alcohol use.

These associated conditions, considered broadly, may have significant implications for HIV-infected veterans, including shortened survival. It is also important to note that these data suggest that the effects of a history of alcohol abuse and current hazardous use seem to be additive. Figure 3 is an example of how alcohol affects, as measured in the VACS, affect both psychiatric and medical outcomes. The role of alcohol in determining outcomes among HIV-infected veterans is multifaceted and must be better defined. Additionally, these preliminary data suggest that HIV providers need to routinely screen and counsel patients for alcohol use as part of standard care. Providers also need to screen and perform interventions on HIV-infected veterans, particularly those who are homeless, have another comorbid disease, or are co-infected with HCV. Alcohol-related interventions need to target reduced alcohol consumption and treatment for alcohol use, particularly hazardous drinking and alcohol dependence.

The VACS study, with its large and diverse cohort of patients and its significantly improved alcohol measures, will provide an opportunity to characterize the role of alcohol use and HIV infection in patient outcomes and to develop effective clinical interventions. VACS will provide critical information needed to inform effective and efficient design of large-scale, multisite intervention studies to decrease the comprehensive harmful effects of alcohol consumption on HIV outcomes. These effects likely go well beyond the known effects of alcohol on adherence (Cook et al., 2001) and risk behavior (Avins et al., 1994; Boscarino et al., 1995; Brown, 2000; Donovan and McEwan, 1995; Hines and Caetano, 1998; Leigh and Stall, 1993; Scheidt and Windle, 1995; Zenilman et al., 1994). These data will also complement data emerging from other NIAAA-supported smaller-scale alcohol intervention studies focused on decreasing HIV risk behavior and improving antiretroviral adherence.

REFERENCES

Avins AL, Woods WJ, Lindan CP (1994) HIV infection and risk behavior among heterosexuals in alcohol treatment programs. JAMA 271:515–518.

Babor TF, de la Fuente JR, Saunders J, Grant M (1992) *AUDIT—The Alcohol Use Disorders Identification Test: Guidelines for Use in Primary Health Care*. World Health Organization, Geneva.

Bangsberg D, Tulsky JP, Hecht FM, Moss AR (1997) Protease inhibitors in the homeless. JAMA 278:63–65.

Bica I, McGovern B, Dhar R, Stone D, McGowan K, Scheib R, Snydman DR (2001) Increasing mortality due to end-stage liver disease in patients with human immunodeficiency virus infection. Clin Infect Dis 32:492–497.

Boscarino JA, Avins AL, Woods WJ, Lindan CP, Hudes ES, Clark W (1995) Alcohol-related risk factors associated with HIV infection among patients entering alcoholism treatment: implications for prevention. J Stud Alcohol 56:642–653.

Brown EJ (2000) AIDS-related risk behavior of young college students. ABNF J 37–41.

Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA (1998) The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Arch Intern Med 158:1789–1795.

Chesney MA (2000) Factors affecting adherence to antiretroviral therapy. Clin Infect Dis (Suppl 2) 30:S171–S176.

Chesney MA, Ickovics JR, Chambers DB, Gifford AL, Neidig J, Zwickl B, Wu AW (2000) 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 12:255–266.

Cohen MH, French AL, Benning L, Kovacs A, Anastos K, Young M, Minkoff H, Hessol NA (2002) Causes of death among women with human immunodeficiency virus infection in the era of combination antiretroviral therapy. Am J Med 113:91–98.

Conigliaro J, Gordon AJ, McGinnis KA, Rabeneck L, Justice AC (2003) How harmful is hazardous alcohol use and abuse in HIV infection: do providers know who is at risk? J Acquir Immune Defic Syndr 33:521–525.

Conigliaro J, Lofgren RP, Hanusa BH (1998) Screening for problem drinking: impact on physician behavior and patient drinking habits. J Gen Intern Med 13:251–256.

Conigrave KM, Hall WD, Saunders JB (1995) The AUDIT questionnaire: choosing a cut-off score. Addiction 90:1349–1356.

Cook RL, Sereika SM, Hunt SC, Woodward WC, Erlen JA, Conigliaro J (2001) Problem drinking and medication adherence among persons with HIV infection. J Gen Intern Med 16:83–88.

Cottler LB, Robins LN, Helzer JE (1989) The reliability of the CIDI-SAM: a comprehensive substance abuse interview. Br J Addict 84:801–814.

Donovan C, McEwan R (1995) A review of the literature examining the relationship between alcohol use and HIV-related sexual risk-taking in young people. Addiction 90:319–328.

Fiellin DA, Reid MC, O'Connor PG (2000) Outpatient management of patients with alcohol problems. Ann Intern Med 133:815–827.

Gordon AJ, Maisto SA, McNeil M, Kraemer K, Conigliaro RL, Kelley ME, Conigliaro J (2001) Three questions can detect hazardous drinkers. J Fam Pract 4:313–320.

Haubrich R, Thompson M, Schooley R, Lang W, Stein A, Sereni D, et al. (1999a) A phase II safety and efficacy study of amprenavir in combination with zidovudine and lamivudine in HIV-infected patients with limited antiretroviral experience. Amprenavir PROAB2002 Study Team. AIDS 13:2411–2420.

Haubrich RH, Little SJ, Currier JS, Forthal DN, Kemper CA, Beall GN, et al. (1999b) The value of patient-reported adherence to antiretroviral therapy in predicting virologic and immunologic response. California Collaborative Treatment Group. AIDS 13:1099–1107.

Hines AM, Caetano R (1998) Alcohol and AIDS-related sexual behavior among Hispanics: acculturation and gender differences. AIDS Educ Prev 10:533–547.

Isaacson JH, Butler R, Zacharek M, Tzelepis A (1994) Screening with the Alcohol Use Disorders Identification Test (AUDIT) in an inner-city population. J Gen Intern Med 9:550–553.

Justice AC, Holmes W, Gifford AL, Rabeneck L, Zackin R, Sinclair G, et al. (2001) Development and validation of a self-completed HIV symptom index. J Clin Epidemiol (Suppl 1) 54:S77–S87.

Justice AC, Wagner J, Fusco GP, et al. (2002) HIV survival: liver function tests independently predict survival. 14th International AIDS Conference, Barcelona, Spain.

Kalichman SC, Rompa D, Cage M (2000) Sexually transmitted infections among HIV seropositive men and women. Sex Transm Infect 76:350–354.

Khan KN, Yatsuhashi H (2000) Effect of alcohol consumption on the progression of hepatitis C virus infection and risk of hepatocellular carcinoma in Japanese patients. Alcohol Alcohol 35:286–295.

Klein MB, Lalonde RG, Suissa S (2003) The impact of hepatitis C virus coinfection on HIV progression before and after highly active antiretroviral therapy. J Acquir Immune Defic Syndr 33:365–372.

Leigh BC, Stall R (1993) Substance use and risky sexual behavior for exposure to HIV. Issues in methodology, interpretation, and prevention. Am Psychol 48:1035–1045.

Loguercio C, Di Pierro M, Di Marino MP, Federico A, Disalvo D, Crafa E, Tuccillo C, Baldi F, del Vecchio Blanco C (2000) Drinking habits of subjects with hepatitis C virus-related chronic liver disease: prevalence and effect on clinical, virological and pathological aspects. Alcohol Alcohol 35:296–301.

Miller WR, Tonigan JS, Longabaugh R (1995) *The Drinker Inventory of Consequences (DrInC): An Instrument for Assessing Adverse Consequences of Alcohol Abuse—Test Manual (NIAAA Project MATCH Monograph Series)*. Government Printing Office, Washington, DC.

Monga HK, Rodriguez-Barradas MC, Breaux K, Khattak K, Troisi CL, Velez M, Yoffe B (2001) Hepatitis C virus infection-related morbidity and mortality among patients with human immunodeficiency virus infection. Clin Infect Dis 33:240–247.

National Institutes of Health Consensus Development Conference Statement: management of hepatitis C (2002) Gastroenterology 123:2082–2099.

Paterson DL, Swindells S, Mohr J, Brester M, Vergis EN, Squier C, Wagener MM, Singh N (2000) Adherence to protease inhibitor therapy and outcomes in patients with HIV infection. Ann Intern Med 133:21–30.

Poynard T, Bedossa P, Opolon P (1997) Natural history of liver fibrosis progression in patients with chronic hepatitis C. Lancet 349:825–832.

Quan CM, Krajden M, Grigoriew GA, Salit IE (1993) Hepatitis C virus infection in patients infected with the human immunodeficiency virus. Clin Infect Dis 17:117–119.

Saunders JB, Conigrave KM (1990) Early identification of alcohol problems. CMAJ 143:1060–1069.

Scheidt DM, Windle M (1995) The alcoholics in treatment HIV risk (ATRISK) study: gender, ethnic and geographic group comparisons. J Stud Alcohol 56:300–308.

Shapiro MF, Morton SC, McCaffrey DF, Senterfitt JW, Fleishman JA, Perlman JF, et al. (1999) Variations in the care of HIV-infected adults in the United States: results from the HIV Cost and Services Utilization Study. JAMA 281:2305–2315.

Sherman KE, Rouster SD, Chung RT, Rajicic N (2002) Hepatitis C virus prevalence among patients infected with human immunodeficiency virus: a cross-sectional analysis of the US adult AIDS Clinical Trials Group. Clin Infect Dis 34:831–837.

Skinner HA, Horn JL (1984) *Alcohol Dependence Scale (ADS) User's Guide*, pp 1–35. Addiction Research Foundation, Toronto.

Skinner HA, Sheu W (1982) Reliability of alcohol use indices. The Lifetime Drinking History and the MAST. J Stud Alcohol 43:1157–1170.

Smith MY, Rapkin BD, Winkel G, Springer C, Chhabra R, Feldman IS (2000) Housing status and health care service utilization among low-income persons with HIV/AIDS. J Gen Intern Med 15:731–738.

Sobell LC, Sobell MB (1994) *Timeline Followback (TLFB): User's Manual*. Addiction Research Foundation, Toronto, Ontario.

Song JY, Safaeian M, Strathdee SA, Vlahov D, Celentano DD (2000) The prevalence of homelessness among injection drug users with and without HIV infection. J Urban Health 77:678–687.

Staples CT Jr, Rimland D, Dudas D (1999) Hepatitis C in the HIV (human immunodeficiency virus) Atlanta V.A. (Veterans Affairs Medical Center) Cohort Study (HAVACS): the effect of coinfection on survival. Clin Infect Dis 29:150–154.

Stein MD, Rich JD, Maksad J, Chen MH, Hu P, Sobota M, et al. (2000) Adherence to antiretroviral therapy among HIV-infected methadone patients: effect of ongoing illicit drug use. Am J Drug Alcohol Abuse 26:195–205.

Wagner JH, Justice AC, Chesney M, Sinclair G, Weissman S, Rodriguez-Barradas M (2001) Patient- and provider-reported adherence: towards a clinically useful approach to measuring anti-retroviral adherence. J Clin Epidemiol 54:S91–S98.

Zenilman JM, Hook EW, Shepherd M, Smith P, Rompalo AM, Celentano DD (1994) Alcohol and other substance use in STD clinic patients: relationship with STDs and prevalent HIV infection. Sex Transm Dis 21:220–225.

The American Journal on Addictions 13:33–45, 2004

ISSN: 1055-0496 print / 1521-0391 online
DOI: 10.1080/10550490490265307

Burden of Medical Illness in Drug- and Alcohol-dependent Persons Without Primary Care

Israel De Alba, M.D., M.P.H., Jeffrey H. Samet, M.D., M.A., M.P.H., Richard Saitz, M.D., M.P.H.

Little is known about the frequency, severity, and risk factors for disease in drug- and alcohol-dependent persons without primary medical care. Our aims are to assess the burden of medical illness, identify patient and substance dependence characteristics associated with worse physical health, and compare measures of illness burden in this population. This was accomplished through a cross-sectional study among alcohol-, heroin- or cocaine-dependent persons without primary medical care who were admitted to an urban inpatient detoxification unit. The mean age of these patients was 35.7 (SD 7.8) years; 76% were male and 46% were Black. Forty-five percent reported being diagnosed with a chronic illness, and 80% had prior medical hospitalizations. The mean age-adjusted SF-36 Physical Component Summary (PCS) score was lower than the general U.S. population norm (44.1 vs 50.1; $p < 0.001$). In multivariable analysis, female gender (adjusted mean change in PCS score: -3.71 points, $p = .002$), problem use of hallucinogens (-3.51, $p = 0.013$), heroin (-2.94, $p = 0.008$), other opiates (-3.20, $p = .045$), living alone (-3.15, $p = .023$), having medical insurance (-2.26, $p = 0.014$) and older age ($-.22$ points per year, $p = 0.001$) were associated with worse health. From these data, it seems that alcohol- and drug-dependent persons without primary medical care have a substantial burden of medical illness compared to age- and gender-matched U.S. population controls. While the optimal measure of medical illness burden in this population is unclear, a variety of health measures document this medical illness burden in addicted persons. (Am J Addict 2004;13:33–45)

Received September 20, 2001; revised January 7, 2002; accepted May 7, 2002.
From the Division of General Internal Medicine and Primary Care, University of California, Irvine, Calif. (Dr. De Alba); the Section of General Internal Medicine, Department of Medicine, Boston University School of Medicine, Boston, Mass. (Drs. Samet and Saitz); and the Departments of Epidemiology (Dr. Saitz) and Social and Behavioral Sciences (Dr. Samet), Boston University School of Public Health, Boston, Mass.
Presented in part at the College on Problems of Drug Dependence 62nd annual scientific meeting, San Juan, PR, June 19, 2000, and the National Research Service Award annual meeting, Los Angeles, CA, June 24, 2000.
Address correspondence to Dr. De Alba, 100 Theory, Suite 110, Irvine, CA 92697. E-mail: idealba@uci.edu.

Substance abuse and dependence are prevalent, costly, and have a significant negative impact on both individual and public health. An estimated 8.2 million Americans are dependent on alcohol and 3.6 million on illicit drugs.[1] Substance abuse is related to almost one third of all newly diagnosed AIDS cases[2] and 555,000 annual emergency department visits.[3] Each year, about 100,000 deaths in the United States are related to alcohol consumption and 12,000 to illicit drug abuse.[4] At least 72 conditions requiring hospitalization are partially or wholly attributable to substance abuse.[5] For example, alcohol abuse can result in liver disease, cardiomyopathy, hypertension, and upper digestive tract malignancies, among others.[6] Illicit drug use is a major risk factor for infective endocarditis, chronic viral hepatitis, and AIDS.[7]

Despite the many documented substance abuse-related medical conditions and the high societal cost, many drug- and alcohol-dependent persons lack primary medical care. Continuity of care could yield opportunities for preventive, diagnostic, and therapeutic interventions.[8-10] Evidence suggests that access to primary care has potential to improve not only general health but also drug abuse-related outcomes.[11-15] Despite the opportunities for benefit, nearly half of all patients presenting for publicly supported substance abuse treatment in one urban center did not have an established primary care relationship.[16] Emergency departments are often the only source of care for many persons with addictions.[17,18] These settings rarely have adequate linkages with substance abuse treatment and primary medical care.[19]

Although psychiatric comorbidities and trauma have been well-described, the burden of medical illness in substance-dependent persons has been minimally examined. Little is known about the frequency, severity, and risk factors for disease in this sizeable group. Better understanding of these issues could inform health care providers and policy makers seeking to enhance access to more effective health care for this high risk, vulnerable population. Thus, we sought to assess the burden of medical illness, identify patient and substance dependence characteristics associated with worse physical health, and compare measures of disease burden in drug- and alcohol-dependent persons without primary medical care.

METHODS

Study Design and Population

We performed a cross-sectional analysis of data from patients enrolled in the Health Evaluation and Linkage to Primary care (HELP) study. The HELP study is a randomized, controlled trial assessing the effectiveness of a multidisciplinary clinic for linking patients in a residential detoxification program to primary medical care. All patients in the detoxification unit without a primary care provider were randomized to the intervention, a clinical evaluation in the HELP clinic by a nurse, social worker, and physician, and facilitated referral to an off-site primary care clinic or to standard care. Enrollment took place at an urban inpatient detoxification unit from June 1997 to March 1999. All admissions were screened for eligibility. The Institutional Review Board at Boston Medical Center approved the study.

All patients enrolled in the randomized, controlled trial were included in this cross-sectional study. Eligible patients were eighteen years of age or older and reported alcohol, heroin, or cocaine as first or second drugs of choice. Exclusion criteria were as follows: having a primary care provider and having seen that provider at least on one occasion in the past two years, pregnancy, Mini-Mental State examination score less than 21, a lack of

fluency in either English or Spanish, fewer than three contacts available to facilitate follow-up, or specific plans to leave the Boston area within the next 24 months.

Data Collection

Trained research associates interviewed patients in the detoxification unit 24–48 hours post-admission. The interview included a 60–90 minute questionnaire on sociodemographics and medical and substance abuse history. The questionnaire was translated into Spanish, back translated into English, and checked for accuracy.

Independent Variables

Sociodemographic characteristics, problem substances, and substance abuse severity were the variables of interest. Sociodemographics included age, gender, race/ethnicity, primary language (English, Spanish, or other), place of birth (United States or other), income, education, marital status, living alone in the past six months, and occupation. We also recorded employment pattern and health insurance in the previous six months. Homelessness was defined as spending one or more nights in an overnight shelter or on the street in the previous six months. Annual income was categorized into the following three groups, according to maximum earnings in any of the previous five years: low income (less than $20,000), intermediate income ($20,000 to $50,000), and higher income (more than $50,000).

Problem drug use was defined as any illicit drug use, or alcohol to intoxication, three or more times per week for a year or more or five times in the previous month. Substance dependence diagnoses per se were not determined in this sample of patients admitted for inpatient detoxification. Substances considered were: alcohol, cocaine, heroin, opiates other than heroin, cannabis, amphetamines, hallucinogens, inhalants, barbiturates, and sedatives. Addiction severity was assessed using the Addiction Severity Index (ASI)[20] and the Alcohol Dependence Scale (ADS).[21] The ADS has a range from zero to 47, with higher values indicating more dependence. Variables of interest were the ASI alcohol and drug composite scores, years of problem substance use, average number of drinks per day in the past month, and injection drug use. The ASI-alcohol and ASI-drug scores have a range from zero to one, with larger values indicating greater severity of use.

Dependent Variables

The primary dependent variable of interest was burden of medical illness assessed by the Physical Component Summary (PCS) score and physical health scale scores from the Short Form Health Survey (SF-36).[22] The SF-36 was designed for use in clinical practice and research and include individual multi-item scales that assess eight health concepts:

1. limitations in physical activities because of health problems
2. limitations in social activities because of physical or emotional problems
3. limitations in usual role activities because of physical health problems
4. bodily pain
5. general mental health (psychological distress and well-being)
6. limitations in usual role activities because of emotional problems
7. vitality (energy and fatigue)
8. general health perceptions.[22]

The PCS is a summary measure considering only the SF-36 components that are related to physical health. We also used the medical composite score from the Addiction Severity Index (ASI-medical)[23] and self-report of one or more physician-diagnosed chronic conditions as

additional burden of medical illness measures. The ASI-medical score has a range from zero (no need for treatment) to one (greatest severity of medical problems).

Statistical Analysis

All analyses were performed using SAS statistical software (SAS Institute, Cary, NC). The data analysis was done in four phases. First, descriptive statistics were generated for each study variable. Second, mean PCS scores in study participants adjusted for the age distribution in the U.S. population were compared to the general U.S. population stratified by gender.[22] Third, to identify risk factors for increased burden of illness, we performed bivariate and multivariable analyses. Chi-square tests were used to compare categorical variables and t-tests or Pearson correlation analyses (or Spearman rank tests, as appropriate) for continuous variables. In a secondary analysis, we compared the frequency of problem alcohol and heroin use, ASI-medical scores, and frequency of medical conditions between men and women. Three multivariable regression models were constructed, one for each dependent variable. We used linear regression for the PCS and ASI-medical scores and logistic regression for the presence of chronic conditions. Variables significant at $p \leqslant 0.2$ in bivariate analysis or considered to be clinically relevant were entered into the regression models. Fourth, the three burden of medical illness measures were compared using Pearson correlation analyses.

RESULTS

Demographics

Of a total of 2,062 screened patients admitted for detoxification, 1,420 (68%) were excluded because of an established primary care relationship ($n = 978$), living outside the Boston area ($n = 204$), having fewer than three contacts available ($n = 91$) or a Mini-Mental State score of < 21 ($n = 58$), pregnancy ($n = 51$), or other reasons ($n = 38$). Twenty-seven percent (172/642) of eligible subjects refused to participate.

The study included 470 individuals without primary medical care (see Table 1). The mean age was 35.7 years (*SD* 7.8). English was the first language for 89% of participants, and Spanish was for 8%. Eighty seven percent were born in the United States. While only 8% were married, most had children (70%) and lived with someone (78%).

Eighty-two percent of subjects had problem use of more than one substance (see Table 2). Alcohol problem use was the most common; more men than women reported problem alcohol drinking (90% vs 74%, $p < 0.001$). The mean ASI-alcohol score was 0.5 (*SD* 0.3), the ADS (Alcohol Dependence Scale) score was 16.8 (*SD* 1.8), and the average number of drinks per day in the month prior to detoxification was 21 (*SD* 20). Sixty-five percent had problem cocaine use and 38% problem heroin use. There was no gender difference in frequency of problem heroin use (men, 38%; women, 40%; $p = 0.731$). The mean ASI-drug score was 0.2 (*SD* 0.1), and 36% reported a history of injection drug use.

Burden of Medical Illness

The mean PCS (Physical Component Summary) score in study subjects was 48.1 (*SD* 10.75). After adjusting for age, the mean score for study participants was on average 6 points lower than the mean for the general U.S. population (44.1 vs 50, $p < 0.001$). The mean and age-adjusted PCS scores for men in the study were 49.1 and 44.9, respectively (see Figure 1). Female subjects had a mean and age-adjusted PCS score of 44.9 and 41.4, respectively (see Figure 2).

TABLE 1. **Sociodemographic Characteristics of Drug- and Alcohol-dependent Persons without Primary Medical Care in a Detoxification Unit ($n=470$)**

	n	%
Age (years)		
18–24	28	6
25–34	195	41
35–44	181	38
45–54	55	12
> 55	11	2
Gender		
Male	357	76
Female	113	24
Race/ethnicity		
Black	216	46
White	174	37
Hispanic	52	11
Other	28	6
Education		
≥12 years	324	69
Employment		
Full time	193	41
Part time	80	17
Unemployed	179	38
Other	19	4
Homeless	221	47
Any health insurance	146	31
Current or past smoking	418	89
Highest annual income*		
< $20,000	268	57
$20,000–$50,000	155	33
> $50,000	47	10

*in any of the previous 5 years.

All individual SF-36 scale scores except physical functioning were lower in study subjects compared to the general U.S. population. Among the physical health scale scores, physical role functioning was the most affected; study participants scored 34.5 points lower on average than the general population (see Figure 3).

The mean ASI-medical score in study subjects was .35 (*SD* .35). There was a trend towards lower scores in women (0.34 vs. 0.40, $p=0.135$). Eighty percent of study participants reported one or more previous hospitalizations for medical problems (mean 3.9 times per patient), 21% were prescribed medications for a chronic medical illness, and 61% had experienced medical problems in the last thirty days. More than half reported being bothered by these medical problems (54%) and that treatment for the problems was important (55%).

Forty-seven percent of study participants reported having at least one chronic medical condition, and 20% reported two or more. Asthma/chronic obstructive

TABLE 2. **Frequency and Duration of Problem Substance Use ($n = 470$)**

	n	%	Problem use duration (years)
Alcohol	397	85	14
Cannabis	323	69	10
Cocaine	304	65	9
Heroin	177	38	9
Hallucinogens	99	21	3
Sedatives	86	18	7
Opiates, other than heroin	69	15	6
Amphetamines	59	13	4
Barbiturates	44	9	6
Inhalants	14	3	2
>1 of the above	385	82	11

pulmonary disease and high blood pressure were the most common, and over one-third of patients (35%) had an episodic medical condition in the past six months (see Table 3). Almost one-half of participants (47%) reported having had at least one sexually transmitted disease in the past, and 4% had one in the previous 6 months; 72% reported previous HIV testing, and 2.6% received a positive result. There were no

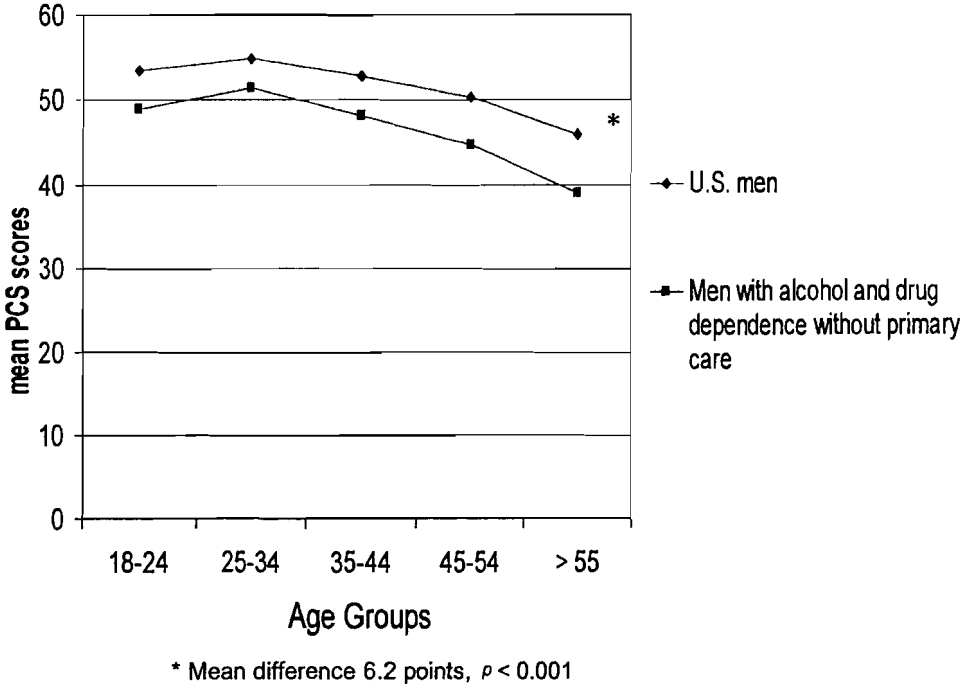


FIGURE 1. *PCS scores for men in the U.S. population and in men with alcohol and drug dependence without primary medical care.*

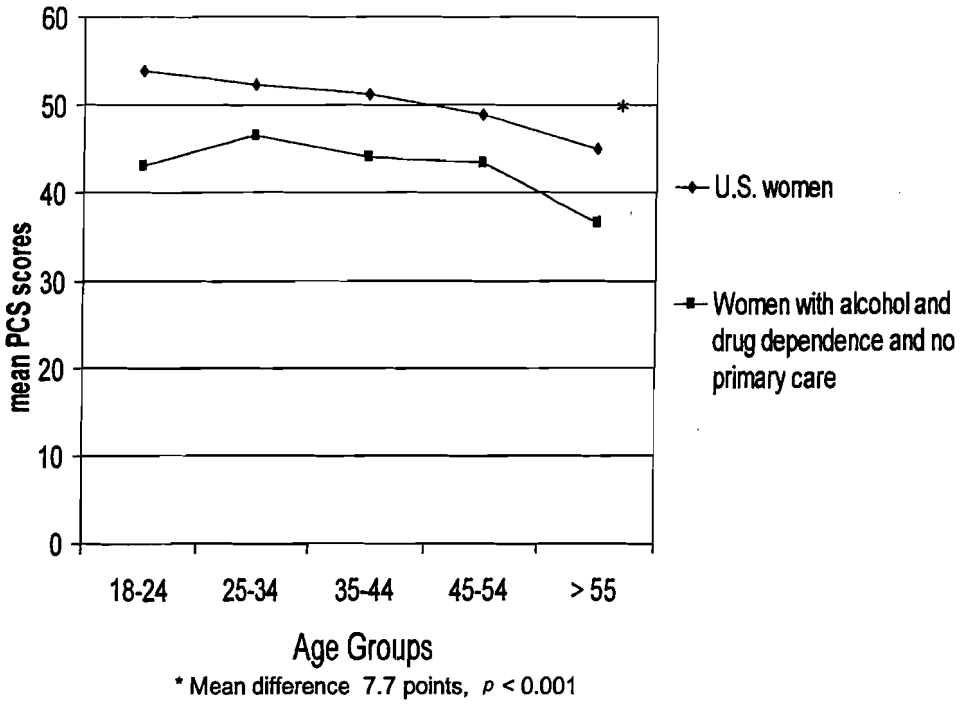


FIGURE 2. PCS scores for women in the U.S. population and in women with alcohol and drug dependence without primary medical care.

gender differences in frequency of chronic (men 45% vs women 53%, $p = 0.151$) or episodic (men 34% vs women 40%, $p = 0.189$) medical conditions in the past six months, but women had a significantly higher incidence of sexually transmitted diseases as compared to men (9% vs 2.5%, $p = 0.002$).

Bivariate Analyses

The following sociodemographic variables were significantly associated with worse physical health status (lower PCS scores) in bivariate analyses: older age, unemployment, female gender, living alone, having medical insurance, and low income (data not shown). Problem use of heroin and other opiates, barbiturates, hallucinogens, and cannabis, as well as addiction severity reflected by the ADS, the ASI-alcohol score, number of drinks in the previous month, and ever using injection drugs, were also significantly associated with worse physical health status.

Multivariable Analyses

As shown in Table 4, the following variables were associated with worse physical health (lower PCS scores) in a multivariable analysis: female gender (adjusted mean change in PCS score: −3.71 points), problem use of hallucinogens (−3.51), heroin (−2.94), or other opiates (−3.20), living alone (−3.15), having

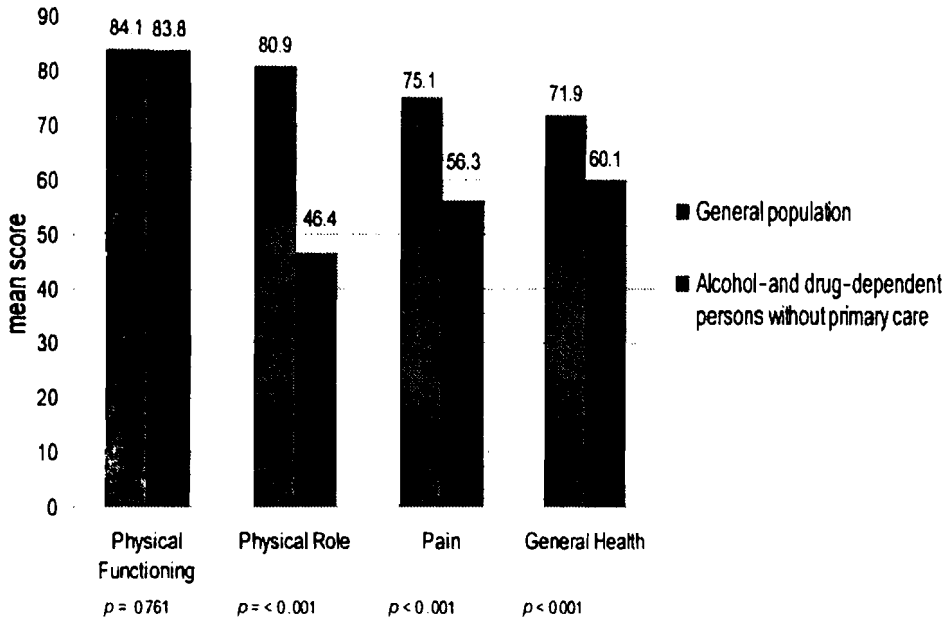


FIGURE 3. Physical health SF-36 scale scores in the general U.S. population and in drug- and alcohol-dependent persons without primary medical care.

medical insurance (−2.26), and older age (−0.22 points per year). Independent risk factors for worse health as measured by the other two measures, ASI-medical score and a presence of a chronic condition, differed from risk factors for worse health as measured by the PCS. Problem alcohol use and having medical insurance were risk factors for worse physical health as assessed by the ASI-medical score. Race/ethnicity other than White, Black, or Hispanic; having medical insurance; and older age increased the risk of having a chronic condition.

Comparison Among Burden of Medical Illness Measures

We found a moderate negative correlation between the PCS and ASI-medical scores ($r = -0.54$, $p = 0.001$), but low correlations between the PCS scores and presence of a chronic condition ($r = -0.29$, $p = 0.001$) and between the ASI-medical score and presence of chronic condition ($r = 0.33$, $p = 0.001$).

DISCUSSION

Drug- and alcohol-dependent persons without primary care have a significant burden of medical illness. In this young sample, most had previous hospitalizations for medical problems, almost half reported having a chronic medical condition, and one in three reported an episodic medical illness in the past six months. Physical health status, as measured by the Physical Component Summary (PCS) of the Short Form health survey, was worse than in the general U.S. population.

Several factors were associated with worse physical health in this population. We found significantly lower PCS scores among women; persons with problem use of heroin, hallucinogens, and other opiates;

TABLE 3. **Chronic, Episodic, and Sexually Transmitted Diseases in Alcohol- and Drug-dependent Persons without Primary Care ($n = 470$)***

Chronic conditions	*n*	%
Asthma/chronic obstructive pulmonary disease	94	20
Hypertension	75	16
Chronic liver disease	56	12
Seizures/epilepsy	52	11
Chronic arthritis/osteoarthritis	38	8
Peripheral neuropathy	14	3
Myocardial infarction	14	3
HIV/AIDS	13	3
Heart conditions other than coronary artery disease	9	2
Diabetes mellitus	8	2
Cancer, any type	6	1
Congestive heart failure	6	1
Stroke	4	1
Any	219	47
Episodic conditions, past six months		
Low back pain	48	10
Hematemesis	38	8
Skin abscesses/infections	30	6
Hepatitis	28	6
Peptic ulcer	23	5
Abdominal pain†	21	4
Chest pain while on cocaine†	18	4
Pneumonia	15	3
Pancreatitis	7	1
Jaundice	5	1
Any	164	35
Sexually Transmitted Diseases, ever		
Gonorrhea	122	26
Chlamydia	56	12
Syphilis	33	7
Genital warts	19	4
Genital herpes	14	3
Any‡	221	47

*Self-report.
†Requiring emergency department visit or overnight hospital stay.
‡All sexually transmitted diseases listed and HIV/AIDS.

older individuals; and those who lived alone, had medical insurance, or were unemployed. The gender difference was not explained by an increased frequency of medical conditions or by greater use of substances that could be associated with an increased burden of illness. The slightly higher incidence in sexually transmitted

TABLE 4. **Multivariable Analyses: Factors Associated with Increased Burden of Illness in Drug- and Alcohol-dependent Persons without Primary Care**

Factors associated with worse PCS scores	Adjusted mean change In PCS	*p*
Female gender	−3.71	0.002
Hallucinogens	−3.51	0.013
Problem use of opiates other than heroin	−3.20	0.045
Living alone	−3.15	0.023
Problem heroin use	−2.94	0.008
Having medical insurance	−2.38	0.014
Unemployment	−2.26	0.026
Age (year)	−0.22	0.001
Factors associated with worse ASI-medical scores	**Adjusted mean change In ASI-medical**	***p***
Problem alcohol use	0.10	0.032
Having medical insurance	0.07	0.031
Factors associated with presence of a chronic condition	**Adjusted mean change Odds ratio**	**95% CI**
Other race*	2.53	1.01–6.35
Having medical insurance	1.95	1.29–2.92
Age (year)	1.05	1.01–1.07

*races other than White, Black, or Hispanic.

diseases reported among women is unlikely to explain the difference in burden. The magnitude of the difference in PCS scores between this sample and the general population was greater than the decline caused by many chronic conditions such as angina, chronic lung disease, diabetes mellitus, and myocardial infarction.[22] The annual incidence rate of pneumonia and prevalence rates of peptic ulcer, hypertension, diabetes mellitus, and stroke among study subjects are more than double the rates found in similar age strata in the general population.[24] The incidence rates for syphilis, gonorrhea, and chlamydia in the previous six months among study participants are more than double the annual rates reported in civilian populations by state health departments.[25]

Differential impact of substance abuse by gender has been documented particularly with regard to alcohol. As compared to men, women in previous studies had higher risk of developing cirrhosis, alcoholic hepatitis,[26] cardiomyopathy, and myopathy.[27] Our findings suggest that regardless of the drug of choice, women with substance dependence suffer from worse physical health. As is the case with alcohol, susceptibility to the adverse health effects of other drugs may be greater in women than in men. Heroin and opiates have been associated with many medical complications and unhealthy lifestyle choices,[7] but the effect of hallucinogens on physical health status is less well known. We do not know if specific types of hallucinogens had a greater impact than others on burden of illness since hallucinogen types were not asked. Social factors like living alone and being unemployed have been reported to be associ-

ated with an increased burden of medical illness[28,29] The fact that having medical insurance was associated with increased burden may sound counterintuitive, but having insurance may have made subjects more likely to know about and therefore report the presence of medical conditions. Alternatively, worse health may have led subjects to be more likely to have health insurance. Disability would make them eligible for Medicaid. More frequent presentation to health care may have increased the likelihood that institutions would facilitate enrollment in public health plans. We did not find an association between drug of choice and having health insurance to suggest differential impact of employability by drug of choice. Older age is known to be associated with worse physical health.

The measures chosen in this study to assess burden of medical illness had only low to moderate correlation with one another. The association was moderate between the PCS and ASI-medical scores, but both correlated poorly with the presence of a chronic condition. Some studies using SF-36-based instruments have shown a detrimental effect of substance abuse on physical health scores,[30-32] but others have shown little or no impact.[33-35] For instance, in a cross-sectional study among 1,333 primary care patients, alcohol use frequency, quantity, abuse, or dependence did not affect PCS scores, and only dependence affected SF-36 scale scores.[36] Among crack cocaine users, the baseline SF-36 scores had no decline during a two-year period despite continued drug use. Additionally, the physical role, vitality, general health, and emotional role scale scores were not affected by frequency of crack use.[37] Whether the lack of impact on health in these two cases truly represents no impact on health or is due to low sensitivity of the scale in this population is not clear. In our study, the PCS was able to detect a significant decrease in health. Seemingly, the PCS, the ASI-medical score, and the presence of a chronic condition assess illness burden from different perspectives and may be complementary. It is not clear which scale is best suited to assess medical illness burden in persons with alcohol or drug dependence. Our findings suggest that each may have value but that one alone may be inadequate. Measures to assess burden of illness in this specific population need to be further developed.

Several limitations of the present study should be considered. The inferences about the causal relationship between risk factor and increased burden of medical illness are limited by the cross-sectional design of the study. Chronic conditions were assessed by self-report but had to be physician diagnosed. Interviews administered shortly after admission to detoxification, when patients were in physical and mental distress, might adversely impact PCS and ASI-medical scores, but subjects were interviewed only after acute symptoms subsided. Subjects were not characterized in terms of substance use disorder diagnoses using *DSM-IV* criteria, which may make comparisons difficult with other studies. However, given the clinical setting, reasons for admission, and results on other measures, all were likely substance dependent. Lastly, the results presented might not be applicable to persons not included in the study, like pregnant women, persons with dementia or not in a detoxification unit, and users whose drug of choice is not alcohol, cocaine, or heroin. But the severity of medical illness among our sample as reflected by ASI-medical scores was comparable to other drug-dependent persons at admission to detoxification programs.[30]

Drug- and alcohol-dependent patients are a vulnerable underserved population with a significant burden of medical illness. Given the potential for effective interventions to decrease this burden (ie, screening and early treatment of sexually

transmitted diseases, tuberculosis, and HIV; hepatitis and pneumococcal vaccines), enhanced efforts should be made to increase linkage to comprehensive and longitudinal primary medical care. The results of our study help to characterize the magnitude of medical burden and identify persons at greatest risk for worse physical health among persons with alcohol and drug dependence without primary medical care; they may also inform the design of improved linkage strategies and medical services aimed at this population.

Funding for this work was provided by National Research Service Award T32 PE1002802 (Dr. De Alba) and by grants R01-10870 from the National Institute of Alcohol Abuse and Alcoholism, Rockville, Md., R-01-10019 from the National Institute on Drug Abuse, Bethesda, Md. (Drs. Saitz and Samet), and M01 RR00533 from the Boston University School of Medicine General Clinical Research Center for the National Center for Research Resources, Bethesda, Md. (Dr. Aram Chobanian).

REFERENCES

1. National Household Survey on Drug Abuse 1999. *Substance abuse and mental health services administration.* Available at: http://www.samhsa.gov/oas/. Accessed November 11, 2000.
2. Center for Disease Control and Prevention. *HIV/AIDS Surveillance Supplemental Report.* 1999;5(3):4–7.
3. Substance Abuse and Mental Health Services Administration. Office of Applied Studies. Drug Abuse Warning Network. *Detailed Emergency Department Tables 1999.* Available at: http://www.samhsa.gov. Accessed December 1, 2000.
4. McGinnis JM, Foege WH. Actual causes of death in the United States. *JAMA.* 1993;270:2207–2212.
5. D'Onofrio G, Bernstein E, Bernstein J, et al. Patients with alcohol problems in the Emergency Department; part I: improving detection. SAEM Substance Abuse Task Force. Society for Academic Emergency Medicine. *Acad Emerg Med.* 1998;5:1200–1209.
6. National Institute on Alcohol Abuse and Alcoholism. *Ninth Report to the U.S. Congress on Alcohol and Health from the Secretary of Health and Human Services.* NIH Publication No 97–4017. Rockville, MD: NIH; June 1997.
7. Cherubin CE, Sapira JD. The medical complications of drug addiction and the medical assessment of the intravenous drug user: 25 years later. *Ann Intern Med.* 1993;119: 1017–1028.
8. Donaldson M, Yordy K, Vanselow N. *Defining Primary Care: An Interim Report.* Washington DC: National Academy Press 1994.
9. Samet JH, Saitz R, Larson MJ. A case for enhanced linkage of substance abusers to primary medical care. *Substance Abuse.* 1996; 17:181–192.
10. Shea S, Misra D, Ehrlich MH, Field L, Francis CK. Predisposing factors for severe, uncontrolled hypertension in an inner-city minority population. *N Engl J Med.* 1992;327:776–781.
11. Samet JH, Friedmann PD, Saitz R. Benefits of linking primary medical care and substance abuse services: patient, provider, and societal perspectives. *Arch Intern Med.* 2001;161:85–91.
12. Ettner SL. The relationship between continuity of care and the health behaviors of patients: does having a usual physician make a difference? *Med Care.* 1999;37:547–555.
13. 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 long term study with randomized controls. *Alcoholism Clin Exp Res.* 1983;7: 203–209.
14. Selwyn PA, Feingold AR, Lezza A, et al. Primary care for patients with immunodeficiency virus (HIV) in a methadone maintenance program. *Ann Intern Med.* 1989;110:761–763.
15. O'Connor PG, Molde S, Henry SP, Shockor WT, Schottenfeld RS. Human immunodeficiency virus infection in intravenous drug

users: a model for primary care. *Am J Med.* 1992;93:382–386.
16. 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.
17. Fox K, Merrill JC, Chang H, Califano JA. Estimating the cost of substance abuse to the Medicaid Hospital Care Program. *Am J Public Health.* 1995;85:48–54.
18. Cherpitel CJ. Alcohol and injuries resulting from violence: a review of emergency room studies. *Addiction.* 1994;89:157–165.
19. Bernstein E, Tracey A, Bernstein J, Williams C. Emergency department detection and referral rate for patients with problem drinking. *Substance Abuse.* 1996;17:69–76.
20. McLellan AT, Luborsky L, Cassiola J, et al. New data from the Addiction Severity Index: reliability and validity in three centers. *J Nerv Mental Dis.* 1985;173:412–423.
21. Skinner HA, Allen BA. Alcohol dependence syndrome: measurement and validation. *J Abnormal Psych.* 1982;91:199–209.
22. Ware JE, Kosinski M, Keller SD. SF-36 Physical and Mental Health Summary scales: a user's manual. Boston, MA: The Heath Institute, New England Medical Center; December 1994. Second printing (revised).
23. McLellan AT, Kushner H, Metzger D, et al. The fifth edition of the Addiction Severity Index. *J Subst Abuse Treat.* 1992;9:199–213.
24. Adams PF, Hendershot GE, Marano MA. Current estimates from the National Health Interview Survey, 1996. National Center for Health Statistics. *Vital Health Stat.* 1999;10:12–82.
25. National Center for Health Statistics. *Health, United States, 1996–97.* Hyattsville, Md.: Public Health; 1998:173.
26. Becker U, Deis A, Sorensen TI, et al. Prediction of risk of liver disease by alcohol intake, sex and age: a prospective population study. *Hepatology.* 1996;23:1025–1029.
27. Urbano-Marquez A, Estruch R, Fernandez-Sola J, Nicolas JM, Pare JC, Rubin E. The greater risk of alcoholic cardiomyopathy and myopathy in women compared to men. *JAMA.* 1995;271:149–154.
28. Jin RL, Shah CP, Svoboda TJ. The impact of unemployment on health: a review of the evidence. *CMAJ.* 1995;153:529–540.
29. Seeman TE. Social ties and health: the benefits of social integration. *Ann Epidemiol.* 1996;6:442–451.
30. Stein MD, Mulvey KP, Plough A, Samet JH. The functioning and well being of persons who seek treatment for drug and alcohol use. *Substance Abuse.* 1998;10:75–84.
31. Ryan FC, White JM. Health status at entry to methadone maintenance treatment using the SF-36 health survey questionnaire. *Addiction.* 1996;91:39–45.
32. Daeppen JB, Krieg MA, Burnand B, Yersin B. MOS-SF-36 in evaluating health-related quality of life in alcohol-dependent patients. *Am J Alcohol Drug Abuse.* 1998;24:685–694.
33. Friedmann PD, Jin L, Karrison T, et al. The effects of alcohol abuse on the health status of older adults seen in the emergency department. *Am J Drug Alcohol Abuse.* 1999;25:529–542.
34. Romeis JC, Waterman B, Scherrer JF, et al. The impact of sociodemographics, comorbidity and symptom recency on health-related quality of life in alcoholics. *J Stud Alcohol.* 1999; 60:652–662.
35. Spitzer RL, Kroenke K, Linzer M, et al. Health-related quality of life in primary care patients with mental disorders. *JAMA.* 1995;274:1511–1517.
36. Volk RJ, Cantor SB, Steinbauer JR, Cass AR. Alcohol use disorders, consumption patterns, and health-related quality of life of primary care patients. *Alcohol Clin Exp Res.* 1997;21:899–905.
37. Falck RS, Wang J, Siegal HA, Carlson RG. Longitudinal applications of the Medical Outcomes Study 36-item Short-Form Health Survey with not-in-treatment crack-cocaine users. *Med Care.* 2000;38:902–910.

A Pilot Trial of a Telecommunications System in Sleep Apnea Management

Deborah A. DeMolles, BA, David Sparrow, DSc,*† Daniel J. Gottlieb, MD, MPH,*† and Robert Friedman, MD‡*

Background: Continuous positive airway pressure (CPAP) is an effective therapy for obstructive sleep apnea syndrome (OSAS), although many patients have difficulty adhering to this therapy. The purpose of this study was to investigate the effectiveness of totally automated telephone technology in improving adherence to prescribed CPAP therapy.

Research Design: This pilot study was a randomized clinical trial in 30 patients being started on CPAP therapy for OSAS. Patients were randomly assigned to use of a computer telephone system designed to improve CPAP adherence (telephone-linked communications for CPAP [TLC-CPAP]) in addition to usual care (n = 15) or to usual care alone (n = 15) for a period of 2 months. TLC-CPAP is a computer-based system that monitors patients' self-reported behavior and provides education and reinforcement through a structured dialogue.

Measures: A sleep symptoms checklist and the Functional Outcomes of Sleep Questionnaire were administered at study entry and at 2-month follow up. Hours of CPAP use at effective mask pressure were measured by the CPAP device, stored in its memory, and retrieved at the 2-month visit.

Results: At 2 months, patients randomized to TLC-CPAP had fewer reported sleep-related symptoms (9.4 vs. 13.4, $P = 0.047$) than those receiving usual care. The average nightly CPAP use in the TLC-CPAP group was 4.4 hours compared with 2.9 hours ($P = 0.076$) in the usual-care group.

Conclusions: This pilot study suggests that patients with OSAS started on CPAP and a concurrently administered automated education and counseling system had better CPAP adherence and better control of OSAS symptoms.

Key Words: OSAS, CPAP therapy, adherence, quality of life

(*Med Care* 2004;42: 764–769)

Obstructive sleep apnea syndrome (OSAS) has been observed in 4% and 2%, respectively, of middle-aged men and women in the United States.[1] Affected persons experience substantial neurobehavioral morbidity, including excessive daytime sleepiness, decreased vigilance, increased auto accidents, impaired quality of life (QOL), and depression. In addition, OSAS has been implicated as a risk factor for hypertension,[2–5] cardiovascular disease,[6,7] and all-cause mortality.[8,9] The initial treatment of choice for most patients with OSAS is nasal continuous positive airway pressure (CPAP) applied by mask during sleep. This therapy has been shown to reduce the frequency of apnea, hypopnea, and oxygen desaturation during sleep, thereby reducing sleep disruption.[10,11] These effects lead in turn to reduced daytime sleepiness and improved vigilance and QOL.[11–13]

Unfortunately, patient adherence to the prescribed use of CPAP is often poor, diminishing the potential benefits of this therapy. Of patients started on nasal CPAP for OSAS, 20% to 33% discontinue CPAP use within 3 to 4 months.[14,15] Even among patients who do not totally discontinue CPAP use soon after it is prescribed, use is less than optimal.[14,16,17]

Patients could discontinue or underuse CPAP therapy for many reasons: inadequate comprehension of the rationale for therapy, including the expected long-term outcomes of CPAP use; side effects; and a lack of knowledge on how to ameliorate these side effects. Patient adherence during the first month after initiation of CPAP therapy is a powerful predictor of long-term adherence,[14,16] and many patients quickly abandon CPAP or start using it only sporadically during the first month. Thus, to enhance long-term adherence, it is crucial to improve patients' understanding of the expected benefits of CPAP use, recognize and address side effects, and monitor and promote adherence to CPAP therapy during the initial period of use.

Side effects of CPAP use are common, although generally minor, and include mask discomfort, dryness of the nose and throat, machine noise, eye irritation, skin redness or ulceration from pressure or allergy, discomfort from applied pressure, claustrophobia, and belching or air swallowing.[15,18,19] Many of these effects, once recognized, are ame-

From the *VA Medical Center, Boston, Massachusetts; †The Pulmonary Center, Boston University School of Medicine, Boston, Massachusetts; ‡Medical Information Systems Unit, Boston Medical Center, Boston University School of Medicine, Boston, Massachusetts.

Supported by the VA Health Services Research and Development Service.

Reprints: Deborah A. DeMolles, BA, VA Medical Center (151N), 150 South Huntington Ave., Boston, MA 02130. E-mail: deborah.demolles@med.va.gov.

ISSN: 0025-7079/04/4208-0764
DOI: 10.1097/01.mlr.0000132353.99209.fe

nable to interventions, including mask adjustment, CPAP pressure adjustment, and humidification.

Telephone-linked communications (TLC) systems offer an effective, low-cost, convenient, and easy-to-use way of providing information, advice, and counseling to improve patient adherence for important health-related behaviors.[20] In previous studies, we have demonstrated 1) the applicability of TLC technology in the care of adults with a wide variety of chronic health conditions[21,22] and 2) its efficacy in modifying patient health behavior (medication-taking, diet, and exercise).[20,21,23]

We designed and constructed a new TLC system for promoting adherence to CPAP that implements and expands on the recommendations of a recent American College of Chest Physicians (ACCP) consensus statement that patient education and monitoring are important, at least during the first month of therapy, to promote long-term adherence to CPAP therapy.[24] We have developed the TLC-CPAP system as an automated, telephone-based, and low-cost intervention modeled on the personnel-intensive intervention of Hoy et al.[25] To obtain preliminary evidence of the feasibility and efficacy of the TLC-CPAP system, we conducted a pilot study of TLC-CPAP among adults with OSAS who were being started on nasal CPAP therapy.

METHODS

Telephone-Linked Communication Technology

The TLC technology is a computer-based telecommunications system that functions as an at-home monitor, educator, and counselor to improve health-related behaviors.[20] TLC speaks to patients over the telephone through computer-controlled digitized human speech. The patients respond by using the touchtone keypad of their telephones. This type of telephonic technology is often called an interactive voice response (IVR) system. TLC asks questions, monitors the patient's self-reported behavior, and provides education and behavioral reinforcement for targeted health-related behaviors. After detailed analysis of actual conversations, TLC conversations have been designed to emulate telephone conversations between patients and health professionals. The words used by TLC to express a particular question or response vary from conversation to conversation to keep the dialogue interesting and as much like human speech as possible.

Like other TLC applications, we designed TLC-CPAP to improve adherence to therapy. The TLC-CPAP content is based on patterns of CPAP adherence[14,16] and side-effect profiles[15,18,19,26] identified in prior studies. The overall outline and sequence of a TLC-CPAP call is shown in Figure 1. For this study, patients were asked to make their first TLC-CPAP call 3 days after starting CPAP therapy (3-day call) and thereafter weekly (1-week call) for a total of 2 months.

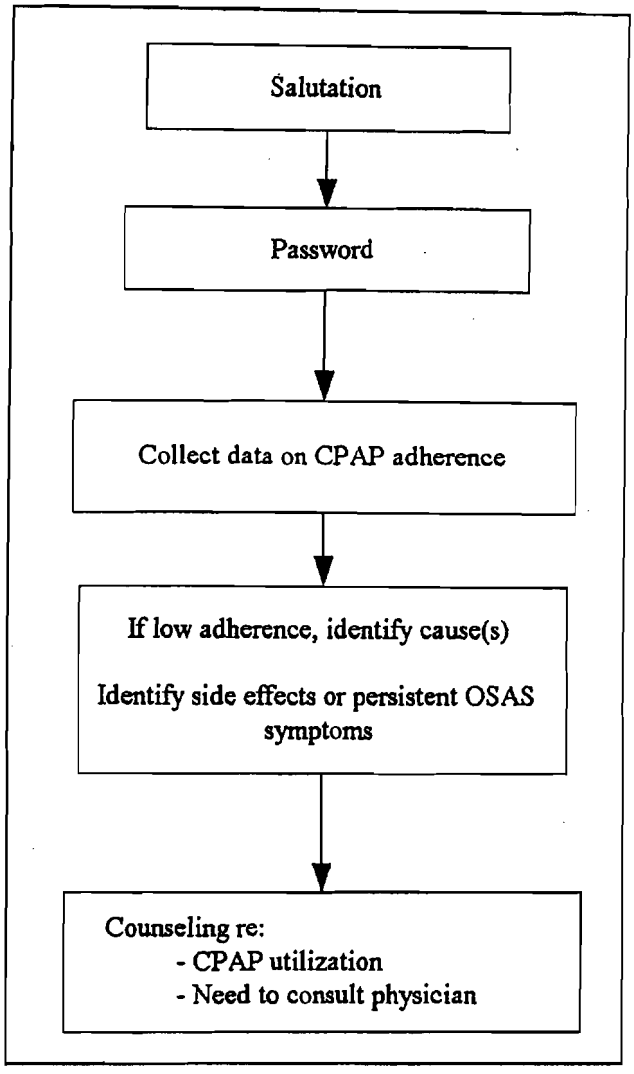


FIGURE 1. Outline of a typical telephone-linked communications for continuous positive airway pressure conversation.

Calls could be made at any time of day that was convenient for the user. If a patient did not make a call to TLC on a scheduled day, TLC called that person the next day, repeating calls periodically during a time period set with the user. If 2 days elapsed from the day of the scheduled call, the system administrator was notified automatically and informed the research assistants working on the project, who then would follow up with the patient to determine why the call was not made.

At the beginning of a call, the user was instructed to enter a personal password, similar to an ATM password, to ensure security and confidentiality. After the confirmed identification of the user, TLC began with an assessment of the frequency and duration of CPAP use during the previous

week (except for the first call, in which 3 days' use were collected). If a patient reported nonuse of the CPAP, or use for fewer than 4 hours per night (on nights they used it) or fewer than 5 nights per week (or fewer than 2 nights in the case of the 3-day call), the system proceeded to ask a series of questions aimed at identifying the cause of CPAP nonadherence (side effects, difficulty using CPAP, lack of perceived benefit, machine malfunction). When CPAP side effects were identified, the severity of each was ascertained. If a patient reported adequate CPAP adherence, the ensuing dialogue reinforced this behavior and then identified the presence of CPAP side effects, if any.

Starting with the first 1-week call, for both CPAP-adherent and -nonadherent patients, the system ascertained the severity of OSAS-related symptoms, including snoring, breathing pauses, and daytime sleepiness. If the patient reported side effects or OSAS symptoms, the system would recommend that the patient contact his or her physician to discuss these problems. At this point in the call, the TLC-CPAP system provided a brief counseling dialogue, focusing on appropriate CPAP use, expected benefits, correct CPAP operating technique, and potential side effects and their treatment. Reinforcement of the need for regular CPAP use was provided at the end of the dialogue (except when a patient had been instructed by a physician to discontinue CPAP), reminding the patient that CPAP use would likely reduce daytime sleepiness and could also have the additional benefit of reducing the risk of cardiovascular disease.

Routine printed reports were sent to the patients' physicians biweekly. These reports included information on the frequency and duration of CPAP use, side effects, and OSAS symptoms. In addition, reports were sent during unscheduled weeks when a patient reported CPAP underuse (for the 3-day call, less than 6 hours over the 3 days; for a 1-week call, less than 4 hours per night on nights the CPAP was used or less than 5 nights per week) or side effects, including mask discomfort, dryness of the nose and throat, claustrophobia, and belching or air swallowing, deemed by the patient to be of moderate-to-severe intensity. Notification was sent to the physician if a patient reported discontinuing CPAP at the physician's instruction to confirm that instruction.

Study Sample

Potential participants were adults starting nasal CPAP therapy through a collaborating home care company (North Atlantic Medical Services, Leominster, MA). To be included in the study, a person had to be aged 18 years or older, English-speaking, have a physician diagnosis of OSAS, and have polysomnography demonstrating >15 episodes of apnea or hypopnea per hour of sleep. Individuals who met these criteria were contacted by a therapist from the home care company by telephone to ask whether they were willing to have a member of the research team speak with them about the study. Patients who agreed were visited by a member of the research team in the person's home along with the therapist initiating CPAP therapy. After the therapist had completed the CPAP setup, including the usual patient education and demonstration of equipment use, the research assistant described the research study and invited the subject to participate. Informed consent was obtained, including consent for release of the sleep study report to the investigators. Individuals were excluded if they reported prior CPAP use. At the conclusion of a baseline examination (see the description of study measures subsequently), eligible participants were randomized to either TLC and usual medical care or usual medical care alone. Subjects assigned to the TLC intervention group were trained to use TLC. A total of 30 subjects were enrolled (15 in each group). All subjects completed a 2-month follow-up evaluation. This research was approved by the Institutional Review Board of the Boston University Medical Center.

Data Collection

Data for analysis were collected during the 2 home visits done 2 months apart. At the baseline examination, height and weight were measured, and subjects completed a checklist of sleep symptoms and the Functional Outcomes of Sleep Questionnaire (FOSQ). At the 2-month examination, the questionnaires were readministered and hours of CPAP use were recorded from a meter.

The FOSQ is a self-reported measure designed to assess the impact of disorders of excessive sleepiness on multiple activities of daily living.[27] The Sleep Symptoms Checklist developed by a research group at the Scottish National Sleep Laboratory[25] measures the frequency of 9 sleep-related symptoms using a 6-point Likert scale, ranging from "never" = 0 to "always" = 5. The Sleep Symptoms Checklist score is calculated by adding the responses for the 9 items (maximal possible score = 45). The average hours per night of CPAP use at effective mask pressure is a primary outcome measure for this study. All devices used in this study were capable of measuring the daily CPAP use at effective mask pressure defined as a pressure within 2 cm of the prescribed CPAP pressure.[14,17] The use-time data stored in the CPAP device memory were retrieved at the 2-month in-home visit.

Data Analysis

All data analysis was performed using SAS 6.12 (SAS Institute, Inc., Cary, NC). Differences between intervention and control groups in the mean values of the outcome measures were analyzed using independent-sample t tests. Analysis of covariance was used to adjust for differences between the groups in baseline characteristics. Because previous research indicates that an adverse impact on CPAP adherence, OSAS symptoms, or QOL resulting from increased monitoring and education is unlikely,[25,28–30] the null

hypothesis of no effect of TLC-CPAP was tested in relation to the directional alternative hypothesis of improvement in these outcome measures.[31]

RESULTS

At baseline, intervention and usual-care subjects had similar characteristics; there were no differences at $P < 0.05$ level (Table 1).

During the 2 months after study entry, the average nightly use of CPAP (averaged across all nights in the study period) was 4.4 hours in the TLC-CPAP intervention group and 2.9 hours in the usual-care group (Table 2). At baseline, intervention and usual-care patients had similar Sleep Symptoms Checklist scores and disease-specific QOL as measured by the FOSQ. After 2 months, TLC-CPAP patients had lower Sleep Symptom Checklist scores (9.4 vs. 13.4; $P = 0.047$) (with lower scores indicating fewer symptoms) than usual-care patients, although there remained little difference in the FOSQ scores (Table 2). Adjustment for baseline characteristics (Table 1) using analysis of covariance had little effect on the results of Table 2.

The 15 patients in the TLC-CPAP group participated in the trial for a mean of 9.2 weeks. Given the prescribed calling plan, we would have expected patients to make an average of 9.3 calls during the trial, but in fact they made an average of 7.9 calls to the system (85%). Of the 7.9 calls made per patient, on average 2.7 calls were made late with 1.5 calls being made after receiving an automated reminder call from the system, whereas another 1.2 calls required a research assistant to follow up after allowing 2 automated reminder calls to be placed.

DISCUSSION

These pilot study results show promising differences between the intervention and usual-care groups, although the small sample size precludes statistical certainty. The use of our automated telephone monitoring system by OSAS patients started on CPAP is associated with increased CPAP use and reduced Sleep Symptoms Checklist score. The average nightly CPAP use was 1.5 hours greater in the intervention group than in the usual-care group for the duration of the study. Although this difference was not statistically significant in this small, pilot study, an effect of this magnitude would be of potentially great clinical significance, because it reflects an approximately 30% increase in average nightly CPAP use over that reported in most clinical trials of CPAP. In other studies, the average nightly duration of CPAP use is significantly correlated with degree of improvement in OSAS symptoms, daytime sleepiness, and general health-related quality of life.[13,32] At follow up, the TLC-CPAP patients' Sleep Symptoms Checklist score was 4 points lower than that of the usual-care group, indicating fewer sleep-related symptoms. There was no significant difference between groups in disease-specific QOL as measured by the FOSQ.

The TLC technology offers a low-cost and easy-to-use means of promoting adherence to CPAP. The rationale for frequent, early monitoring of CPAP use is based on the observation that by 1 month after initiation of therapy, patterns of CPAP use are firmly established.[14,16] Indeed, recent evidence suggests that a pattern of regular CPAP use could be established as early as the first week of therapy and that discomfort from the CPAP device is an important predictor of irregular use.[26] Our TLC-CPAP intervention was modeled on the intervention of Hoy et al.[25] These investigators conducted a randomized, controlled trial of an intensive support intervention among 80 patients with OSAS using CPAP for the first time. The intervention included initial education, an extra 2-night CPAP titration protocol (subsequent to the initial CPAP titration study) to adjust pressure settings, and periodic home visits, mostly during the initial month, by a specially trained nurse. CPAP use was measured by pressure-time meters installed in the CPAP machines. Patients in the intervention group had higher average nightly use of CPAP over 6 months than did control patients (5.4 ± 0.3 vs. 3.8 ± 0.4

TABLE 1. Baseline Characteristics of Study Subjects*

	Intervention (TLC-CPAP)	Control
Age (years)	49.8 ± 15.7	42.0 ± 13.0
Body mass index (kg/m^2)	38.0 ± 12.2	38.1 ± 7.0
Apnea–hypopnea Index	41.8 ± 38.1	38.1 ± 40.1
Sleep Symptoms Checklist	23.7 ± 7.0	25.3 ± 9.1
Functional Outcomes of Sleep Questionnaire	15.3 ± 3.5	13.8 ± 4.6

*All results are mean ± standard deviation. There were no differences between the 2 groups on these characteristics at $P < 0.05$.

TLC-CPAP indicates telephone-linked communications for continuous positive airway pressure.

TABLE 2. Outcome Measures for Study*

	Intervention (TLC-CPAP)	Control	P Value
Average nightly CPAP use (hr) over 2 months	4.4 ± 3.0	2.9 ± 2.4	0.076
Sleep Symptoms Checklist	9.4 ± 6.0	13.4 ± 6.6	0.047
Functional Outcomes of Sleep Questionnaire	18.1 ± 2.8	17.0 ± 3.7	0.171

*All results are mean ± standard deviation.

TLC-CPAP indicates telephone-linked communications for continuous positive airway pressure.

hours [mean ± standard error of mean], respectively, $P = 0.003$). Patients in the intervention group also had fewer sleep symptoms at the 6-month follow up (7 ± 1 vs. 10 ± 1) as measured by the Sleep Symptoms Checklist. The intervention also led to improvements in mood and simple unprepared reaction time. Although the authors could not be sure which component(s) of the intervention accounted for the improvements seen, their impression was that the extra 2-night titration added little.

Other studies of interventions to increase CPAP use have provided conflicting results but have had important methodologic limitations. In an uncontrolled retrospective study, Likar et al.[28] observed that a single 2-hour educational session led to an increase in CPAP use as measured by run-time metering among 25 OSAS patients. Fletcher and Luckett[29] performed a randomized, controlled trial of the impact of human telephone support on objective CPAP use among 10 OSAS patients newly started on this therapy. During the initial 3 months of the intervention, the 6 subjects in the intervention group had somewhat greater CPAP use than did the 4 control subjects (6.6 ± 1.8 vs. 5.6 ± 3.9 hours/night [mean ± standard deviation], respectively), although this difference was not statistically significant. After crossover between experimental conditions, no difference was seen after an additional 3 months. Both the very small sample size and a crossover design that would tend to diminish intervention effects if the benefits of the intervention were sustained limit interpretation of this study. Finally, Chervin et al.[30] reported that OSAS patients who received weekly telephone calls or an educational intervention had greater adherence to CPAP than did a usual-care control group. This study, however, had important methodologic limitations, including a substantially lower severity of OSAS and less severe subjective and objective sleepiness at entry in control subjects than in subjects in either intervention group.

In our study, we observed increased CPAP use (1.5 hours/night more in TLC-CPAP users) similar to that in the study of Hoy et al.[25] described previously. The reduction in Sleep Symptoms Checklist score demonstrated in our data is greater than that observed by Hoy et al.[25] Thus, these preliminary data suggest that TLC-CPAP use leads to an increase in CPAP adherence and an improvement in sleep apnea symptoms similar to those of the more personnel-intensive program of Hoy et al.[25] Despite the face validity of the Sleep Symptoms Checklist as a measure of clinical OSAS severity, data relating checklist scores to other measures of clinical outcome are limited. To put these results in a clinical context, using the approach suggested by Cohen for interpreting effect size defined as the mean difference/standard deviation of the difference, the effect size of 0.64 for the Sleep Symptoms Checklist is suggestive of moderate clinical significance.[33] In the study by Hoy et al., the similar improvements in CPAP use and Sleep Symptoms Checklist scores were associated with both subjective improvement in mood and objective improvement in reaction time.[25] Another way to put these results in context is to compare the effect of the TLC-CPAP intervention to the overall effect of CPAP treatment on OSAS symptoms. The effect observed in this study was 27% as great as that observed for CPAP therapy in a group of 62 OSAS patients studied by Kingshott et al., who were similar in age, body mass index, and baseline Sleep Symptoms Checklist score to the subjects of the present study.[32] The effect of the TLC-CPAP intervention on FOSQ score had an effect size of 0.34, consistent with small clinical significance.[33]

In summary, adherence to nasal CPAP prescribed for OSAS is often poor, limiting the health benefits of this expensive therapy. Nurse-administered patient education and monitoring of CPAP use through home visits has been shown to be effective in significantly improving CPAP adherence; however, given the logistic complexity and cost of delivering this service, it is unlikely to be disseminated widely in clinical practice. The current study provides preliminary evidence that the use of advanced telecommunications technology to perform patient education and monitoring provides an acceptable and inexpensive alternative that could produce a similar improvement in CPAP adherence.

REFERENCES

1. Young T, Palta M, Dempsey J, et al. The occurrence of sleep-disordered breathing among middle-aged adults. *N Engl J Med.* 1993;328:1230–1235.
2. Nieto FJ, Young TB, Lind BK, et al. Association of sleep-disordered breathing, sleep apnea, and hypertension in a large community-based study. *JAMA.* 2000;283:1829–1836.
3. Peppard PE, Young T, Palta M, et al. Prospective study of the association between sleep-disordered breathing and hypertension. *N Engl J Med.* 2000;342:1378–1384.
4. Hla KM, Young TB, Bidwell T, et al. Sleep apnea and hypertension: a population-based study. *Ann Intern Med.* 1994;120:382–388.
5. Stradling JR, Crosby JH. Relation between systemic hypertension and sleep hypoxemia or snoring: analysis of 748 men drawn from general practice. *BMJ.* 1990;300:75–78.
6. Waller PC, Bhopal RS. Is snoring a cause of vascular disease? An epidemiologic review. *Lancet.* 1989;1:143–146.
7. Hung J, Whitford EG, Parson RW, et al. Association of sleep apnea with myocardial infarction in men. *Lancet.* 1990;336:261–264.
8. Partinen M, Jamieson A, Guilleminault C. Long-term outcome for obstructive sleep apnea syndrome patients: mortality. *Chest.* 1988;94:1200–1204.
9. He J, Kryger MH, Zorick FJ, et al. Mortality and apnea index in obstructive sleep apnea: experience in 385 male patients. *Chest.* 1988;94:9–14.
10. Sullivan CE, Issa FG, Berthon-Jones M, et al. Reversal of obstructive sleep apnoea by continuous positive airway pressure applied through the nares. *Lancet.* 1981;1:862–865.
11. Rajagopal KR, Bennett LL, Dillard TA, et al. Overnight nasal CPAP improves hypersomnolence in sleep apnea. *Chest.* 1986;90:172–176.
12. Engleman HM, Martin SE, Kingshott RN, et al. Randomized, placebo controlled trial of daytime function after continuous positive airway pressure (CPAP) therapy for the sleep apnoea/hypopnoea syndrome. *Thorax.* 1998;53:341–345.
13. Engleman HM, Kingshott RN, Wraith PK, et al. Randomized placebo-controlled crossover trial of continuous positive airway pressure for mild

sleep apnea/hypopnea syndrome. *Am J Respir Crit Care Med.* 1999;159: 461–467.
14. Reeves-Hoche MK, Meck R, Zwillich CW. Nasal CPAP: an objective evaluation of patient compliance. *Am J Respir Crit Care Med.* 1994; 149:149–154.
15. Rolfe I, Olson LG, Saunders NA. Long-term acceptance of continuous positive airway pressure in obstructive sleep apnea. *Am Rev Respir Dis.* 1991;144:1130–1133.
16. Kribbs NB, Pack AI, Kline LR, et al. Objective measurement of patterns of nasal CPAP use by patients with obstructive sleep apnea. *Am Rev Respir Dis.* 1993;147:887–895.
17. Engleman HM, Martin SE, Douglas NJ. Compliance with CPAP therapy in patients with the sleep apnoea/hypopnoea syndrome. *Thorax.* 1994; 49:263–266.
18. Hoffstein V, Viner S, Mateika S, et al. Treatment of obstructive sleep apnea with nasal continuous positive airway pressure. *Am Rev Respir Dis.* 1992;145:841–845.
19. Waldhorn RF, Herrick TW, Nguyen MC, et al. Long-term compliance with nasal continuous positive airway pressure therapy of obstructive sleep apnea. *Chest.* 1990;97:33–38.
20. Friedman RH. Automated telephone conversations to assess health behavior and deliver behavioral interventions. *J Med Syst.* 1998;22:95–102.
21. Friedman RH, Kazis LE, Jette A, et al. A telecommunications system for monitoring and counseling patients with hypertension: impact on medication adherence and blood pressure control. *Am J Hypertens.* 1996;9: 2852–2892.
22. Young M, Sparrow D, Gottlieb D, et al. A telephone-linked computer system for COPD care. *Chest.* 2001;119:1565–1575.
23. Delichatsios HK, Friedman RH, Glanz K, et al. Randomized trial of a 'talking computer' to improve adults' eating habits. *Am J Health Promot.* 2001;15:215–224.
24. Loube DI, Gay PC, Strohl KP, et al. Indications for positive airway pressure treatment of adult obstructive sleep apnea patients. A consensus statement. *Chest.* 1999;115:863–866.
25. Hoy CJ, Vennelle M, Kingshott RN, et al. Can intensive support improve continuous positive airway pressure use in patients with the sleep apnea/hypopnea syndrome? *Am J Respir Crit Care Med.* 1999; 159:1096–1100.
26. Weaver TE, Kribbs NB, Pack AI, et al. Night-to-night variability in CPAP use over the first three months of treatment. *Sleep.* 1997;20:278–283.
27. Weaver TE, Laizner AM, Evans LK, et al. An instrument to measure functional status outcomes for disorders of excessive sleepiness. *Sleep.* 1997;20:835–843.
28. Likar LL, Panciera TM, Erickson AD, et al. Group education sessions and compliance with nasal CPAP therapy. *Chest.* 1997;111:1273–1277.
29. Fletcher EC, Luckett RA. The effect of positive reinforcement on hourly compliance in nasal continuous positive airway pressure users with obstructive sleep apnea. *Am Rev Respir Dis.* 1991;143:936–941.
30. Chervin RD, Theut S, Bassetti C, et al. Compliance with nasal CPAP can be improved by simple interventions. *Sleep.* 1997;20:284–289.
31. Rosner B. *Fundamentals of Biostatistics.* Belmont, CA: Duxbury Press: 1995.
32. Kingshott RN, Vennelle M, Hoy CJ, et al. Predictors of improvements in daytime function outcomes with CPAP therapy. *Am J Respir Crit Care Med.* 2000;161:866–871.
33. Cohen J. *Statistical Power for the Behavioral Sciences.* Hillsdale, NJ: Erlbaum Press; 1988.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Drug and Alcohol Dependence 73 (2004) 159–166
DRUG and ALCOHOL DEPENDENCE
www.elsevier.com/locate/drugalcdep

Inconsistent condom use among HIV-infected patients with alcohol problems

Vera Ehrenstein [a], Nicholas J. Horton [b], Jeffrey H. Samet [c,*]

[a] *Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Department of Epidemiology, Boston Medical Center, Boston University Schools of Medicine and Public Health, 91 East Concord Street, Suite 200, Boston, MA 02118, USA*
[b] *Department of Mathematics, Smith College, Northampton, MA 01063, USA*
[c] *Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Departments of Medicine, and Social and Behavioral Sciences, Boston Medical Center, Boston University Schools of Medicine and Public Health, 91 East Concord Street, Suite 200, Boston, MA 02118, USA*

Received 13 February 2003; received in revised form 8 September 2003; accepted 9 October 2003

Abstract

Background: Unsafe sexual behavior is common among persons with negative or unknown HIV status and it is augmented by alcohol use in some populations. We examined the association between alcohol consumption level (abstinent, moderate, at-risk) and inconsistent condom use in a cohort of HIV-infected individuals with a history of alcohol problems. *Methods:* Subjects ($n = 345$) had up to seven structured interviews over 36 months. Identical questions on alcohol consumption and inconsistent condom use were asked at each interview. We used generalized estimating equations (GEE) multivariate logistic regression for repeated measurements analysis. We adjusted for potential confounding factors and explored possible interactions. *Results:* At baseline, 132 (38%) participants reported inconsistent condom use. We detected a significant ($P = 0.0002$) interaction between alcohol consumption and injection drug use (IDU) variables. Among active injection drug users, at-risk drinking was associated with inconsistent condom use, adjusted odds ratio (OR; 95% confidence interval) 4.3 (1.5, 12.2). Among those who did not inject drugs, at-risk drinking and inconsistent condom use were not associated, 0.7 (0.4, 1.3). Inconsistent condom use was more common among women, those believing condoms to be 'a hassle', and persons living with a partner. *Conclusion:* In HIV-infected drug-injecting individuals, excessive use of alcohol is associated with unsafe sexual practices.

Keywords: Human immunodeficiency virus; Alcohol; Condom use; Generalized estimating equations

1. Introduction

While the diagnosis of human immunodeficiency virus (HIV) infection may lead to the initiation of either safer sex or sexual abstinence (Pierret, 2000), unsafe sexual behavior remains common among people with HIV, as documented by both self-report and by laboratory tests for sexually-transmitted infections (STI) (Avants et al., 2000; Erbelding et al., 2000; Kalichman, 1999; Kalichman et al., 2000). Recent studies of HIV-infected patients revealed a 12% prevalence of STI in Atlanta (Kalichman et al., 2000), and a 7.5% prevalence of gonorrhea or chlamydia infection in Baltimore (Erbelding et al., 2000).

HIV-related morbidity and mortality have been reduced by the use of antiretroviral therapy (ART) (Palella et al., 1998). At the same time, the availability of treatment helped fuel the belief that undetectable viral load precludes viral transmission, leading to decreased vigilance in adhering to safer sexual practices (Kalichman et al., 2001; Katz et al., 2002). This complacency is an important public health concern, as it promotes the spread of sexually-transmitted infections, including HIV and its resistant strains (Little et al., 2002).

While sharing contaminated injection instruments is an efficient mode of HIV transmission, unprotected sexual intercourse is the most 'universally available' mode, facilitating HIV spread beyond traditional risk groups. Unsafe sexual practices in non-HIV-infected populations have been linked to illicit drug use (Buchacz et al., 2001; Woods et al., 1996, 2000), African–American ethnicity (Buchacz et al., 2001), low socio-economic status (Buchacz et al., 2001),

* Corresponding author. Tel.: +1-617-414-7399; fax: +1-617-414-4676.
E-mail address: jsamet@bu.edu (J.H. Samet).

0376-8716/$ – see front matter © 2003 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.drugalcdep.2003.10.011

younger age (Buchacz et al., 2001), and depressive symptoms (Skrondal et al., 2000).

While not universally supported by published data (Messiah et al., 1998; Weatherburn et al., 1993), the association of alcohol with unsafe sex in HIV-uninfected persons has been described both generally (Caetano and Hines, 1995; McEwan et al., 1992) and in special populations, such as young adults (Bagnall et al., 1990; Graves, 1995; Seage et al., 1998), drug users (Rees et al., 2001; Stein et al., 2000), and bisexual men (Wold et al., 1998). Active alcohol use as a risk factor for risky sexual behavior is suggested by Scheidt (1999), who found that compared with the pre-treatment period of active alcohol use, subjects had fewer sexual partners after substance abuse treatment.

Among those who are infected with HIV, risky sexual behavior can lead to adverse outcomes. On the individual level, it increases the risk of re-infection with a potentially drug resistant strain of HIV. From the public health perspective, unsafe sex facilitates the spread of HIV infection. Although the role of alcohol use in relation to unsafe sex has been studied in non-HIV-infected special populations, its impact in persons with HIV has not been fully explored. Prevalence of HIV infection among alcoholics is higher than that in the general population (Avins et al., 1994; Woods et al., 2000). At the same time, studies report over 40% prevalence of alcohol problems among HIV-infected persons (Lefevre et al., 1995; Samet et al., in press). These findings underscore the importance of understanding the impact of alcohol use on unsafe sexual behaviors in HIV-infected persons, particularly since sexual transmission of HIV is, in principle, preventable. Our principal hypothesis in this study was that, in HIV-infected persons with a history of alcohol problems, high current alcohol consumption is associated with higher risk of inconsistent condom use. In order to better assess alcohol's independent effect on this behavioral outcome, we controlled for other subject characteristics, such as injection drug use (IDU), sexual orientation, partner status, gender, education, and ethnic origin.

2. Methods

2.1. Study design & data collection

Participants ($n = 349$) were recruited between July 1997 and July 2001 into the HIV–Alcohol Longitudinal cohort, a follow-up study of HIV-infected patients with past or current history of alcohol problems. Participants had up to seven interviews, with the baseline visit generally occurring within 2 weeks of initial screening and subsequent interviews scheduled at 6-month intervals. No follow-up occurred after July 2001.

Structured face-to-face confidential interviews were administered in English or Spanish by trained staff. For the Spanish interviews, standardized scales in Spanish were used when available; the remaining portions of the questionnaires were translated from English, back-translated to check for accuracy, and corrected. Laboratory values of HIV RNA and CD4 cell counts measured within 3 months of each visit were obtained from medical records whenever available. If not available, blood samples were drawn during the visit by nursing staff. Most interviews took place at the General Clinical Research Center of Boston University School of Medicine. Participants were compensated US$ 20 or an equivalent gift certificate to a local grocery store. Institutional Review Boards of Boston Medical Center (BMC) and Beth Israel Deaconess Medical Center (BIDMC) approved the study.

2.2. Participants

2.2.1. Recruitment sites

All participants resided in the Greater Boston area. The majority of participants were recruited as they initiated HIV medical care at BMC's HIV Diagnostic Evaluation Unit, a clinic that provided multidisciplinary initial evaluation for HIV-infected persons (Samet et al., 1995). Other subjects were recruited from the following sites: primary care clinics at BMC and BIDMC; a methadone clinic in Boston; a respite facility for homeless persons; referrals by friends; media announcements and flyers placed in HIV/AIDS social service agencies in the Boston area. Persons recruited outside BMC or BIDMC were pre-screened by telephone, and potentially eligible individuals were invited to complete the screening process in person.

2.2.2. Eligibility

The inclusion criteria were: (1) confirmed HIV infection; and (2) a history of alcohol problems as determined either by endorsement of two or more items on the CAGE alcohol screening questionnaire (Ewing, 1984; Mayfield et al., 1974; Samet et al., in press), or by clinical judgment of one of two physician study investigators. The exclusion criteria were: (1) inability to understand spoken English or Spanish; (2) evidence of impaired cognitive function as determined by a low (<21) score on the Mini Mental State Examination (Folstein et al., 1975); or (3) plans to leave the Boston area during the next 2 years. Patients who were not in care at BMC or BIDMC were asked to document their HIV diagnosis by providing either HIV testing documentation or their HIV prescription medications.

2.3. Measures

2.3.1. Variables measured at baseline and at each visit (repeated measures)

2.3.1.1. Dependent variable. Consistency of condom use was assessed using the item from the risk assessment battery (RAB) "In the past 6 months, how often did you use condoms when you had sex?" (Navaline et al., 1994). Inconsistent condom use was defined as less than 100%

reported use in the last 6 months before each visit. Those reporting either 100% condom use or sexual abstinence formed the comparison group.

2.3.1.2. Main independent variable. Alcohol consumption in the 30 days before each assessment was used as a measure of the usual pattern of use. It was calculated using alcohol quantity and frequency questions, and the Addiction Severity Index (ASI) (Leonhard et al., 2000). We classified alcohol use as 'abstinent', 'moderate' and 'at-risk', based on the National Institute on Alcohol Abuse and Alcoholism recommendations, which define moderate drinking as no more than 14 drinks per week for men, and no more than seven drinks per week for women (National Institute on Alcohol Abuse and Alcoholism, 1995). Since we did not measure daily or weekly alcohol use in this study, we computed average weekly consumption from the data on the 30-day reported frequency of use and typical daily amount. Based on these calculations, we assigned the alcohol use categories.

2.3.1.3. Other independent variables. The following HIV risk behavior questions covering the last 6 months before each visit, asked as a part of the RAB, were included in the analysis: dichotomous variables for injection drug use; number of sexual partners (one or fewer versus two or more); selling sex for money or drugs; buying sex for money or drugs; and an opinion regarding condoms being 'a hassle to use'. Use of heroin or cocaine in the last 30 days before each visit was assessed by the ASI and each used as a dichotomous variable. In order to account for participants' physical and mental health, we included CD4 cell count ($\leq$350/μl versus >350/μl), receiving ART, and depressive symptoms measured with the Center of Epidemiologic Studies Depression (CES-D) scale with scores potentially ranging from 0 to 60 (Radloff, 1977). To address the possibility that inconsistent condom use patterns may change over time spent in the study, an indicator variable for each timepoint was also included in the analysis.

2.3.2. Variables measured only at baseline (point measures)

We included the following sociodemographic variables: age in years; gender; race/ethnicity (black, white, Latino, other); high school completion; having lived with a partner most of the time in previous 6 months; sexual orientation (heterosexual, gay/lesbian, bisexual); homelessness (defined as spending at least one night in a shelter or on the street in the previous 6 months); and recency of HIV diagnosis ($\leq$12 months versus >12 months before the baseline interview).

2.4. Statistical analysis

We excluded from all analyses four subjects who reported sexual abstinence at each visit, since they were not at risk for inconsistent condom use during the follow-up. In the unadjusted analyses, we used Fisher's exact test, or Chi-square test for trend for categorical variables, and *t*-test or analysis of variance for continuous variables. In the adjusted analysis, we modeled inconsistent condom use at each visit as a function of baseline variables (e.g., gender and education) and variables measured at that visit (e.g., current substance use). Depending on the number of completed visits, each subject contributed from 1 to 7 observations to the analysis. We estimated regression parameters and modeled within-subject correlation using generalized estimating equation (GEE) logistic regression models (Liang and Zeger, 1986). These models account for same-subject association by estimating the within-subject outcome odds ratios (OR) for each pair of timepoints. Main-effect independent variables in the analysis were included based on earlier findings or clinical importance. In addition, we examined pairwise interactions between alcohol consumption and the following covariates: race/ethnicity; gender; high school completion; visit indicator variables, sexual orientation; number of sexual partners; injection drug use; heroin use; cocaine use; buying sex for money and/or drugs; selling sex for money and/or drugs; and ART use. We included interaction terms with the inclusion significance criterion set at 0.05.

2.5. Follow-up visits

In order to assess the degree of association between the study variables and the amount of follow-up visits we modeled the number of completed visits as a linear function of study variables. Since the interviewing ended at a pre-determined date with late-entering subjects having fewer or no follow-up opportunities, we hypothesized that recruitment date would be the most important predictor of the number of visits. We used indicator variables for 6-month periods of recruitment. Also included in the model examining number of completed visits were age, gender, race/ethnicity, recency of HIV diagnosis, sexual orientation, along with baseline values of ART use, injection drug use, inconsistent condom use, and level of alcohol consumption. We used SAS software (SAS/STAT Software: Changes and Enhancements, Release 8.1, 2000) to perform the analyses.

3. Results

3.1. Main analysis

Baseline characteristics of the 345 subjects are summarized in Table 1: mean age was 40.4 years; 79% were men; 67% were non-white; 14% had recent incarceration history; and 29% were homeless. Median (range) CD4 cell count was 352 (0–1401)/μl; mean (S.D.) CES-D score was 22.3 (12.9). The most commonly reported HIV risk factors were: IDU (58%); men having sex with men (19%); and heterosexual sex (22%).

In the 6 months before the baseline interview, 57% of subjects reported no alcohol use, 24% reported moderate drinking, and 19% reported at-risk drinking. At baseline,

Table 1
Baseline characteristics of HIV-infected persons with a history of alcohol problems ($N = 345$)

Characteristic	*N*	Percentage
Men	273	79
Race/ethnicity		
Black	151	44
White	115	33
Latino	75	22
Other	4	1
Graduated from high school	208	60
Homeless	101	29
Employment		
Full/part-time	104	30
Unemployed	128	37
Disability	113	33
Primary risk for HIV[a]		
Injection drug use	201	58
Men having sex with men	65	19
Heterosexual contact	77	22
Blood transfusion	1	1
Lives with a partner	65	19
Sexual orientation[a]		
Heterosexual	246	72
Gay/lesbian	70	20
Bisexual	28	8
Drug use in past 30 days		
Cocaine	84	24
Heroin	36	10
Alcohol use		
Abstinent	197	57
Moderate	82	24
At-risk	66	19

[a] $N = 344$ due to missing data. Nine hundred and ninety-six observations from 345 subjects.

37/345 (11%) of subjects reported being in a methadone program. Of those, eleven reported moderate, and nine reported at-risk alcohol consumption. Compared with the entire cohort, the highest proportions of subjects with at-risk drinking at baseline were observed among subjects who used cocaine (54%) or heroin (42%), among those who lived with a partner (30%), or bought sex for money and/or drugs (30%). The prevalence of at-risk drinking was higher among African–Americans (22%) than among whites (16%) or Latinos (17%).

At baseline, 99/345 (29%) of the subjects reported sexual abstinence. Consistency of condom use was reported as follows: 'all of the time' by 114/345 (33%) subjects; 'most of the time' by 57/345 (16%) subjects, 'some of the time' by 40/345 (12%) subjects, and 'none of the time' by 35/345 (10%) subjects. Table 2 depicts prevalence of inconsistent condom use by subjects' baseline characteristics. Proportions of inconsistent condom users in the 'abstinent', 'moderate' and 'at-risk' drinking categories were, respectively, 32, 41, and 51% ($P = 0.005$ for trend). Other factors

Table 2
Sociodemographic and clinical baseline characteristics of the 345 HIV-infected persons with a history of alcohol problems and prevalence of inconsistent condom use

Characteristic	Prevalence of inconsistent condom use (%)	*P*-value
Gender		0.1
Men	36	
Women	47	
Race/ethnicity		0.2
Black	39	
White	43	
Latino	29	
Other	25	
Graduated from high school		0.02
Yes	43	
No	31	
Lives with a partner		<0.0001
Yes	61	
No	33	
Sexual orientation		0.01
Heterosexual	34	
Gay/lesbian	47	
Bisexual	57	
Recent HIV+ test		0.04
No	36	
Yes	51	
Level of alcohol use		0.02
Abstinent	32	
Moderate	41	
At-risk	51	
Injected drugs		0.004
No	34	
Yes	52	
Cocaine use, past 30 days		<0.0001
No	32	
Yes	57	
Heroin use, past 30 days		0.001
No	36	
Yes	58	
Number of sexual partners		<0.0001
One	41	
Two or more	66	
Bought sex for drugs/money		0.004
No	36	
Yes	54	
Sold sex for drugs/money		<0.0001
No	34	
Yes	70	
'Condoms are a hassle to use'		0.0004
Strongly disagree/disagree	33	
Strongly agree/agree	55	
Receiving ART		0.3
No	42	
Yes	36	
CD4 count		0.0005
≤350 cells/μl	29	
>350 cells/μl	48	

Table 3
Multivariate regression analysis of characteristics associated with inconsistent condom use among HIV-infected persons with a history of alcohol problems[a]

Characteristic	Adjusted odds ratio	95% confidence interval
No concurrent injection drug use		
Abstinent	1	
Moderate	1.5	0.9–2.3
At-risk	0.7	0.4–1.3
Concurrent injection drug use		
Abstinent	1.7	0.9–3.1
Moderate	1.2	0.6–2.5
At-risk	4.3	1.5–12.2
Race/ethnicity		
Black	0.8	0.5–1.2
White	1	
Latino	0.6	0.3–1.0
Other	–	
Sexual orientation		
Heterosexual	1	
Gay/lesbian	2.4	1.4–4.1
Bisexual	1.6	0.8–3.3
Women	2.0	1.2–3.4
Graduated from high school	1.4	0.9–2.1
Lives with a partner	3.5	2.2–5.7
Recent HIV+ test	1.6	1.0–2.7
Cocaine use, past 30 days	1.4	0.9–2.3
Heroin use, past 30 days	0.9	0.5–1.8
Two or more sexual partners	2.7	1.8–4.0
Bought sex for drugs/money	0.8	0.5–1.3
Sold sex for drugs/money	1.8	1.0–3.5
Agree that condoms are hassle to use	2.5	1.8–3.4
Receiving ART	1.0	0.7–1.4
CD4 count >350 cells/μl	1.4	1.0–1.9
CES-D (one-point increase)	1.0	0.9–1.0

[a] $N = 344$ due to missing data. Nine hundred and ninety-six observations from 345 subjects.

significantly associated with higher prevalence of inconsistent condom use in unadjusted analyses were the following: high school education; living with a partner; being of gay, lesbian or bisexual orientation; testing positive for HIV in the previous 12 months; injection drug use; use of cocaine or heroin; having two or more sexual partners; trading sex for money or drugs; believing condom use to be 'a hassle'; and CD4 count of >350 cells/μl.

Among the 345 subjects, there were a total of 996 interviews over the 4-year study period. Table 3 shows the results of the multivariate analysis. We detected statistically significant interaction of alcohol consumption with concurrent (within the same 6-month period) injection drug use ($\chi^2 = 17.1$, d.f. $= 2$, $P = 0.0002$). Injection drug use among subjects not reporting alcohol consumption at a given visit was associated with an adjusted odds ratio (95% confidence interval (CI)) of 1.7 (0.9, 3.1); and at-risk drinking among non-injectors, with an adjusted OR (95% CI) of 0.7 (0.4, 1.3). However, compared with those who neither drank alcohol nor injected drugs, at-risk drinkers with concurrent drug injection had an adjusted OR (95% CI) of 4.3 (1.5, 12.2) for inconsistent condom use. These results indicate that the combined effect of at-risk drinking and injection drug use on inconsistent condom use was greater than would be expected under the assumption of multiplication of the effects.

Among those with active injection drug use, the adjusted OR (95% CI) was 2.6 (0.9, 7.3) for at-risk drinkers and 0.7 (0.3, 1.6) for moderate drinkers, each compared with abstainers. Among those without active injection drug use, the adjusted OR (95% CI) was 0.7 (0.4, 1.3) for at-risk drinkers and 1.5 (0.9, 2.3) for moderate drinkers.

Other characteristics significantly associated with inconsistent condom use in the adjusted analysis included living with a partner, adjusted OR (95% CI) 3.59 (2.2, 5.7); having ≥2 partners, 2.7 (1.8, 4.0); being gay or lesbian, 2.4 (1.4, 4.1); believing that condom use was 'a hassle', 2.5 (1.8, 3.4); and female gender, 2.0 (1.2, 3.4). In addition, borderline significant increase in odds was observed among those selling sex for drugs and/or money, adjusted OR (95% CI), 1.8 (1.0, 3.5); those who had a positive HIV test within 12 months of the baseline interview, 1.6 (1.0, 2.7); and those with CD4 cell counts of >350/μl, 1.4 (1.0–1.9). No appreciable association of race/ethnicity, ART use, or depressive symptoms with inconsistent condom use was seen in the adjusted analysis.

3.2. Assessment of differential follow-up

The median number of completed visits was 3. Alcohol use was borderline significant in predicting the number of visits ($P = 0.06$), with those who used alcohol at baseline completing more visits on average than non-users. As hypothesized, time of recruitment into the study was the most important predictor of the number of completed visits ($P < 0.0001$) with subjects entering at the beginning of the recruitment completing, on average, three interviews more than subjects entering at the end of the recruitment. In the model examining the number of visits, all predictors combined explained 37% of the variation. In the main multivariable analysis, time of recruitment was not a significant predictor of inconsistent condom use ($P = 0.7$), nor did it materially change the OR estimates for the alcohol use variable, so it was not included in the final model.

4. Discussion

In examining the association between alcohol consumption and inconsistent condom use in a cohort of HIV-infected patients with past or current alcohol problems, we found that at-risk drinking was associated with inconsistent condom use among active injection drug users. These findings expand on our understanding of how alcohol use influences HIV risk behaviors in high-risk populations. We demonstrate that at-risk alcohol consumption remains associated

with unsafe sex risk behavior among injection drug users after they learn about their HIV infection.

The result that at-risk drinking was not associated with inconsistent condom use in the absence of active injection drug use was not consistent with our original hypothesis. This finding does suggest that the link between alcohol use and risky sexual behavior is not universal; rather particular groups seem to be at higher risk. The joint effect of drug use and alcohol has been noted by other researchers. Stein et al. found that at-risk alcohol use was associated with needle-sharing behavior among active injection drug users (Stein et al., 2000), while Rees et al. linked at-risk alcohol consumption with sex risk behaviors among women with a history of heroin or cocaine dependence (Rees et al., 2001). Thus, drug use and at-risk drinking seem to magnify each other's detrimental effects on safe HIV-related behavior, as was also observed in our study.

We found a moderately strong association between inconsistent condom use and living with a partner. Regular partners of HIV patients are at the highest risk of HIV, including resistant strains (Little et al., 2002), as the probability of HIV transmission rises proportionately to the number of sexual encounters. Inconsistent condom use with cohabitating partners may be perceived as less risky than that with occasional partners, but such behavior among HIV-infected persons still holds great risk for both partners. This is particularly concerning, given that selling sex for money or drugs was associated with inconsistent condom use in our study, and some subjects who reported sex trading also reported living with a partner. Such behavior patterns may in part account for the 2.6-fold increase in transmission of multidrug resistant strains of HIV in the US between 1995 and 2000 (Little et al., 2002).

We found that participants with gay or lesbian sexual orientation were less likely than heterosexual individuals to use condoms. Despite reports suggesting that in some areas of the United States men who have sex with men have adopted and sustained safer sexual practices (Donovan and Ross, 2000), a European study documents rising incidence of gonorrhea among men who have sex with men and an eight-fold increase in the odds of HIV seroconversion compared with heterosexual men (Donovan et al., 2000). Such reports are consistent with our findings, suggesting that the favorable behavior changes among this group need bolstering. Additionally, we found that women were less likely than men to use condoms in heterosexual intercourse; this may reflect some women's inability to negotiate condom use (Kral et al., 2001).

Both high CD4 cell count and recent HIV diagnosis were associated with increased risk of inconsistent condom use in our HIV-infected cohort. Individuals with higher CD4 cell counts are often asymptomatic and have fewer medical problems. Recent HIV diagnosis may represent a different issue as recent HIV diagnosis is not routinely associated with high CD4 cell counts (Samet et al., 2001). We speculate that these findings indicate that HIV-infected individuals need time to realize the implications of HIV infection for sexual behavior, particularly if their symptoms are mild or absent.

Contrary to the findings by Buchacz et al. (2001), we found higher, not lower, educational level to be associated with inconsistent condom use, and observed no striking differences with respect to inconsistent condom use among ethnic groups. These different findings may be due to the different character of our HIV-infected cohort, consisting of individuals with a history of alcohol problems, predominantly unmarried, recruited largely from an urban Northeastern USA setting. In contrast, Buchacz et al. studied HIV–serodiscordant couples in California.

This study has several limitations. We relied on self-report in measuring most of the study variables. Sexual risk behaviors are subject to social desirability bias and may be under-reported. Among HIV-infected individuals, inconsistent condom use carries additional risk of infecting another individual. We dichotomized the inconsistent condom use variable for the analysis in order to reduce potential mixing of true consistent users with those who may report generally consistent use but in reality do not use condoms 100% of the time. We acknowledge that this analytic approach will lump together different patterns of inconsistent condom use, however, this threshold seemed most consistent with the HIV prevention goal to eliminate HIV transmission risk. There is, however, no reason to suspect that this behavior would be differentially reported in different groups of participants. We attempted to minimize any bias by making the interview process confidential and private. We also note that subjects did report a substantial amount of non-socially desirable behavior. Additionally, subjects were followed for variable lengths of time, with some of them not completing all interviews. We addressed this possibility by examining predictors of the total number of visits. As hypothesized, time of recruitment into the study was the most important predictor of the number of completed visits, and it was found not to be a confounder in our analysis.

Finally, data from each visit were analyzed cross-sectionally, allowing for inferences about association, rather than causation. Evidence suggests, however, that alcohol use affects sexual practices immediately at a given sexual encounter rather than over the long-term (Seage et al., 1998). We thus deliberately chose the cross-sectional over the longitudinal approach, in order to capture the use of alcohol most proximal in time to measured condom use. Nonetheless, the measurement of alcohol use associated with each sexual act, which would be ideal, was not obtained in this study.

The relation between alcohol consumption and inconsistent condom use among HIV-infected persons with a history of alcohol problems is complicated. First, alcohol consumption at high, rather than moderate, level is associated with inconsistent condom use. Second, this association is manifest predominantly in injection drug users. This suggests that efforts to reduce alcohol use in HIV-infected persons who inject drugs may also result in progress towards the

goal of decreasing the frequency of HIV sexual risk-taking. These findings bolster the argument for routine assessment of alcohol use in HIV-infected patients, particularly in those with a history of injection drug use.

Acknowledgements

This study was supported by the following grants from the National Institute on Alcohol Abuse and Alcoholism (NIAAA) of the NIH: R01-AA13766 (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, USPHS grant M01-RR00533. We wish to thank Dr. Anita Palepu for valuable comments regarding this analysis, Ms. Nicole Tibbetts of DM-Stat, Inc. for data management, Ms. Seville Meli for support of manuscript preparation, and Dr. Kenneth A. Freedberg for excellent suggestions upon review of the manuscript.

References

Avants, S.K., Warburton, L.A., Hawking, K.A., Margolin, A., 2000. Continuation of high-risk behavior by HIV-positive drug users. Treatment implications. J. Subst. Abuse Treat. 19, 15–22.

Avins, A.L., Woods, W.J., Lindan, C.P., Hudes, E.S., Clark, W., Hulley, S.B., 1994. HIV infection and risk behaviors among heterosexuals in alcohol treatment programs. JAMA 271, 515–518.

Bagnall, G., Plant, M., Warwick, W., 1990. Alcohol, drugs and AIDS-related risks: results from a prospective study. AIDS Care 2, 309–317.

Buchacz, K., van der Straten, A., Saul, J., Shiboski, S.C., Gomez, C.A., Padian, N., 2001. Sociodemographic, behavioral, and clinical correlates of inconsistent condom use in HIV–serodiscordant heterosexual couples. J. Acquir. Immune Defic. Syndr. 28, 289–297.

Caetano, R., Hines, A.M., 1995. Alcohol, sexual practices, and risk of AIDS among blacks, Hispanics, and whites. J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 10, 554–561.

Donovan, B., Bodsworth, N.J., Rohrsheim, R., McNulty, A., Tapsall, J.W., 2000. Increasing gonorrhoea reports—not only in London. Lancet 355, 1908.

Donovan, B., Ross, M.W., 2000. Preventing HIV: determinants of sexual behaviour. Lancet 355, 1897–1901.

Erbelding, E.J., Stanton, D., Quinn, T.C., Rompalo, A., 2000. Behavioral and biologic evidence of persistent high-risk behavior in an HIV primary care population. AIDS 14, 297–301.

Ewing, J.A., 1984. Detecting alcoholism. The CAGE questionnaire. JAMA 252, 1905–1907.

Folstein, M.F., Folstein, S.E., McHugh, P.R., 1975. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J. Psychiatr. Res. 12, 189–198.

Graves, K.L., 1995. Risky sexual behavior and alcohol use among young adults: results from a national survey. Am. J. Health Promot. 10, 27–36.

Kalichman, S.C., 1999. Psychological and social correlates of high-risk sexual behaviour among men and women living with HIV/AIDS. AIDS Care 11, 415–427.

Kalichman, S.C., Rompa, D., Austin, J., Luke, W., DiFonzo, K., 2001. Viral load, percieved infectivity, and unprotected intercourse (Letters to the Editor). J. Acquir. Immune Defic. Syndr. 28, 303–305.

Kalichman, S.C., Rompa, D., Cage, M., 2000. Sexually transmitted infections among HIV seropositive men and women. Sex. Transm. Infect. 76, 350–354.

Katz, M.H., Schwarcz, S.K., Kellogg, T.A., Klausner, J.D., Dilley, J.W., Gibson, S., McFarland, W., 2002. Impact of highly active antiretroviral treatment on HIV seroincidence among men who have sex with men: San Francisco. Am. J. Public Health 92, 388–394.

Kral, A.H., Bluthenthal, R.N., Lorvick, J., Gee, L., Bacchetti, P., Edlin, B.R., 2001. Sexual transmission of HIV-1 among injection drug users in San Francisco, USA: risk-factor analysis. Lancet 357, 1397–1401.

Lefevre, F., O'Leary, B., Moran, M., Mossar, M., Yarnold, P.R., Martin, G.J., Glassroth, J., 1995. Alcohol consumption among HIV-infected patients. J. Gen. Intern. Med. 10, 458–460.

Leonhard, C., Mulvey, K., Gastfriend, D.R., Shwartz, M., 2000. The Addiction Severity Index: a field study of internal consistency and validity. J. Subst. Abuse Treat. 18, 129–135.

Liang, K.Y., Zeger, S.L., 1986. Longitudinal data analysis using generalized linear models. Biometrika 73, 13–22.

Little, S.J., Holte, S., Routy, J.P., Daar, E.S., Markowitz, M., Collier, A.C., Koup, R.A., Mellors, J.W., Connick, E., Conway, B., Kilby, M., Wang, L., Whitcomb, J.M., Hellmann, N.S., Richman, D.D., 2002. Antiretroviral-drug resistance among patients recently infected with HIV. N. Engl. J. Med. 347, 385–394.

Mayfield, D., McLeod, G., Hall, P., 1974. The CAGE questionnaire: validation of a new alcoholism screening instrument. Am. J. Psychiatry 131, 1121–1123.

McEwan, R.T., McCallum, A., Bhopal, R.S., Madhok, R., 1992. Sex and the risk of HIV infection: the role of alcohol. Br. J. Addict. 87, 577–584.

Messiah, A., Bloch, J., Blin, P., 1998. Alcohol or drug use and compliance with safer sex guidelines for STD/HIV infection. Results from the French National Survey on Sexual Behavior (ACSF) among heterosexuals. Analyse des Comportements Sexuels en France. Sex. Transm. Dis. 25, 119–124.

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 Res. Hum. Retroviruses 10 (Suppl. 2), S281–283.

National Institute on Alcohol Abuse and Alcoholism, 1995. Physicians' Guide to Helping Patients with Alcohol Problems. NIH Publication 35-3769, http://www.niaaa.nih.gov/publications/physicn-text.htm, accessed 1 April 2002.

Palella Jr, F.J., Delaney, K.M., Moorman, A.C., Loveless, M.O., Fuhrer, J., Satten, G.A., Aschman, D.J., Holmberg, S.D., 1998. Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators. N. Engl. J. Med. 338, 853–860.

Pierret, J., 2000. Everyday life with AIDS/HIV: surveys in the social sciences. Soc. Sci. Med. 50, 1589–1598.

Radloff, L.S., 1977. The CES-D scale: a self-report depression scale for research in the general population. Appl. Psychol. Meas. 1, 385–401.

Rees, V., Saitz, R., Horton, N.J., Samet, J., 2001. Association of alcohol consumption with HIV sex- and drug-risk behaviors among drug users. J. Subst. Abuse Treat. 21, 129–134.

Samet, J.H., Freedberg, K.A., Savetsky, J.B., Sullivan, L.M., Stein, M.D., 2001. Understanding delay to medical care for HIV infection: the long-term non-presenter. AIDS 15, 77–85.

Samet, J.H., Libman, H., LaBelle, C., Steger, K., Lewis, R., Craven, D.E., Freedberg, K.A., 1995. A model clinic for the initial evaluation and establishment of primary care for persons infected with human immunodeficiency virus. Arch. Intern. Med. 155, 1629–1633.

Samet, J.H., 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, in press.

SAS/STAT Software, 2000. Changes and Enhancements, Release 8.1. SAS Institute Inc., Cary, NC.

Scheidt, D.M., 1999. HIV risk behavior among alcoholic inpatients before and after treatment. Addict. Behav. 24, 725–730.

Seage III, G.R., Mayer, K.H., Wold, C., Lenderking, W.R., Goldstein, R., Cai, B., Gross, M., Heeren, T., Hingson, R., 1998. The social context of drinking, drug use, and unsafe sex in the Boston Young Men Study. J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 17, 368–375.

Skrondal, A., Eskild, A., Thorvaldsen, J., 2000. Changes in condom use after HIV diagnosis. Scand. J. Public Health 28, 71–76.

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. J. Subst. Abuse Treat. 18, 359–363.

Weatherburn, P., Davies, P.M., Hickson, F.C., Hunt, A.J., McManus, T.J., Coxon, A.P., 1993. No connection between alcohol use and unsafe sex among gay and bisexual men. AIDS 7, 115–119.

Wold, C., Seage III, G.R., Lenderking, W.R., Mayer, K.H., Cai, B., Heeren, T., Goldstein, R., 1998. Unsafe sex in men who have sex with both men and women. J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 17, 361–367.

Woods, W.J., Avins, A.L., Lindan, C.P., Hudes, E.S., Boscarino, J.A., Clark, W.W., 1996. Predictors of HIV-related risk behaviors among heterosexuals in alcoholism treatment. J. Stud. Alcohol 57, 486–493.

Woods, W.J., Lindan, C.P., Hudes, E.S., Boscarino, J.A., Clark, W.W., Avins, A.L., 2000. HIV infection and risk behaviors in two cross-sectional surveys of heterosexuals in alcoholism treatment. J. Stud. Alcohol 61, 262–266.

Journal of Medical Systems, Vol. 28, No. 2, April 2004 (© 2004)

Testing the Usability of Two Automated Home-Based Patient-Management Systems

Ramesh Farzanfar,[1,2] Joseph Finkelstein,[1] and Robert H. Friedman[1]

To explore to what extent observation and semistructured in-depth interviews provide effective tools for usability testing of two automated home-based systems aimed at monitoring patients' health status at home and improving self-care. Telephone-Linked Care for Diet Adherence in Dyslipidemia (TLC-DietAid) used computer telephony to interact with users and Home Asthma Telemonitoring System (HAT System) used a combination of Personal Digital Assistant (palmtops) and the Internet for similar purposes. Both systems were evaluated in two separate pilot studies. Our pilot studies uncovered "medium-specific" and "content-specific" issues that addressed either the process of the interaction or its content. The results demonstrated that patient-users tended to evaluate each system on the basis of how it fit into everyday life and corresponded to personal preferences. The methodology also allowed the system designers to understand users' concerns and the context of adoption in order to introduce necessary changes to the design to address such concerns.

KEY WORDS: evaluation; usability testing; interactive health management systems; human factor.

INTRODUCTION

The use of communication and information technology in health care delivery has had a significant impact on the dynamics of the relationship between health care providers and patients.(1) Nowadays, patients and consumers can receive a plethora of health information and advice through various sources from the Internet to their telephones.(2,3) Some of these systems monitor the health status of patients and notify their heath care providers about important changes in status.(4)These systems have the potential of improving health outcomes, reducing health care costs, and reducing the burden of disease management for patients, caregivers and health care providers, particularly for patients with chronic health conditions.(5–7)The range of services these systems offer run the gamut of simple information to complex monitoring and decision support.(8)

[1]Medical Information Systems Unit, Boston University Medical Campus, 560 Harrison Avenue, Suite 405, Boston, Massachusetts 02118.
[2]To whom correspondence should be addressed; e-mail: rfarazanf@bu.edu.

0148-5598/04/0400-0143/0 © 2004 Plenum Publishing Corporation

The potential of these systems, however, can only be realized with effective utilization by patients. This means that not only the technology should be user-friendly, but also the program design should be tailored to particular needs of patients and thus perceived as helpful and acceptable by them. Therefore, these systems need to be evaluated for usability, helpfulness, and satisfaction from the patients' perspective, as well as their effectiveness in promoting health and controlling disease. Evaluation, as a critical step in the development process of these systems, will require an understanding of the patient-users' opinions and perceptions, preferences, likes and dislikes, etc.

These objectives, however, are not easily fulfilled through controlled, formalized, or structured evaluation methodologies. For example, certain types of usability evaluation methods such as cognitive walkthroughs[9–11] or heuristic analysis[12–14] that are typically performed by the developers or expert evaluators without the user's participation are not exactly suitable for evaluating systems that provide health services to users. Or, more traditional research methodologies[15,16] such as structured questionnaires, even though useful, provide only limited information on usability and acceptability.[17] For such systems, users' active participation and involvement in the evaluation process is a critical factor. Such participation may involve a sample of potential users preferably from the beginning of the system's design process[18] or after a system is built with system improvement as the basic objective.[13] Methods for such evaluations may include using workshops and focus groups[19] as well as ethnographic and contextual inquiry.[20–22]

Ethnography, a method that requires participant observation and immersion in the life-world of study participants[23] is perhaps the most effective method for exploring or evaluating human–computer interaction.[24] However, an ethnographic evaluation is costly, long, and labor intensive as it requires a lengthy period of participant observation and exploration of the subjects' every day lives. Because of these constraints, ethnographic methodologies are not frequently used in the evaluation of health information systems. An appropriate alternative, however, is a qualitative methodology that relies on participant observation and in-depth interviews of shorter durations. Thus, instead of conducting a retrospective, single-study in-depth interview to evaluate a system after its use, observation and in-depth interviews may be conducted as users use a system.[17] We utilized such a methodology in two separate pilot studies to evaluate two different interactive patient care-management systems. Below, we will describe how observation, sufficient training, and semistructured in-depth interviews helped to effectively explore patients' opinions and views on two different types of automated interactive systems: TLC-DietAid (Telephone-Linked Care for Diet Adherence Intervention in Dyslipidemia)[8,pp. 95–101] and HAT System (Home Asthma Telemonitoring System).[25] TLC-DietAid is a totally telephone-based conversation system that educates, advises, and counsels patients with dyslipidemia to modify their dietary behaviors. The HAT system, on the other hand, is based on handheld computing devices in the home linked to a Web-based monitoring and reporting system. The HAT system promotes home-based asthma self-care management.

It was anticipated that the data derived from observation and in-depth interviews conducted with patient-users would enable the investigators to identify major problems experienced by the patients while using these systems in their homes. These

methods would help them implement changes in the systems to make them more usable, helpful, and acceptable to patient-users.

TLC-DietAid System

TLC-DietAid is a computer-based telecommunications system that uses computer telephony to carry out totally automated telephone conversations with patient-users.[6] TLC-DietAid monitors dietary behavior of dyslipidemia (cholesterolemia) patients in their homes and provides them with nutrition information, advice, and behavioral counseling, principally to lower saturated fat in their diets and thus lower low-density lipoprotein cholesterol (LDL-C) in their sera.[8] LDL-C is a major risk factor for coronary heart disease, the most significant cause of death in the United States and other industrialized nations.

During weekly conversations, TLC-DietAid speaks to patients over the telephone using computer-controlled digitized human speech. The patients, in turn, communicate with TLC-DietAid using the touch-tone keypad of their telephones. In usual practice, patients call TLC-DietAid once a week, however, for this evaluation study they used it every day for 2 weeks. Conversations typically last between 5 and 10 min, depending on the number or complexity of the topics addressed and the user's responses. After each conversation, TLC stores the information recorded by patients in a database, for use in tailoring future conversations.

Using the National Cholesterol Education Project (NCEP) Step I guidelines as its basis,[26] TLC-DietAid conversations were designed to modify the users' intake of certain foods: (1) to lower the consumption of whole fat dairy foods and red and processed meat, (2) to increase the consumption of fruits and vegetables, and (3) to increase the consumption of whole grains. For individual users, TLC-DietAid targeted those food groups the person consumed in "unhealthy" amounts, either because they consumed too few fruits and/or vegetables, too much whole fat dairy foods and/or red/processed meat, or too little whole grain dairy foods. The system broadly targeted important nutrients in the diet, such as saturated fat and fiber, both by targeting the relevant food groups (e.g., whole fat dairy and red/processed meats for their saturated fat content), but also by encouraging related dietary behaviors such as trimming fat off meat, avoiding fried foods, etc.

To achieve these goals, the design of TLC-DietAid emulated the clinical strategies used by expert clinicians and the principles of a theoretical model of health behavior change: Social Cognitive Theory (SCT). SCT synthesizes concepts and processes from cognitive, behavioral, and emotional models of behavioral change, and postulates that behavior change is determined by the reciprocal determined nature of person, behavior, and environment. Within this encompassing theory, there are many constructs, four of which are especially relevant to the design of TLC-DietAid, namely behavioral capability, self-efficacy, outcome expectation, and reinforcement.[27]

The Home Asthma Telemonitoring System

The Home Asthma Telemonitoring (HAT) system was designed to address problems in asthma self-care management using a combination of mobile computing and

Internet technologies[25] HAT fully implements the guidelines of National Asthma Education and Prevention Program (NAEPP)[28] and uses state-of-the-art knowledge about the educational, behavioral, cognitive, and organizational components of asthma self-management.[29] The HAT system aims to provide patients with continuous individualized help in the daily routine of asthma self-care, coupled with monitoring of their disease status and ongoing communication with their health care providers. It is designed to detect and provide help to asthma patients who have difficulties in following their self-care plan and to notify health care providers if certain clinical conditions occur to enable timely intervention.

The HAT system consists of a patient unit, a decision support server, and a clinical station. The patient unit includes a Personal Digital Assistant (PDA) and an electronic spirometer. Patients enter answers to asthma clinical status questions on the PDA and perform lung function testing using the spirometer on a regular basis. The spirometer transmits the test results via a serial interface to the PDA. The patient can also use the PDA to transmit a personal message to medical personnel in addition to transmitting their answers to the clinical status questions. Immediately after the completion of the self-testing, all data are sent by the PDA to a remote clinical information server which stores the data in a database which is accessible by the patient's asthma provider at a predesignated secure Internet site. The data can be sent over a standard telephone line or over a wireless network. Patient data for the last 4 months are also stored in the PDA and are available for the patient's review. During each self-testing session, the patient receives feedback messages generated automatically, or sent by medical personnel, via the same system. The HAT decision support server constantly monitors patient self-testing data and alerts the staff if the data indicate that the patient is not self-testing properly. The system also alerts clinicians if any clinically negative trend in the patient's condition occurs indicated by the spirometer or clinical data. This helps clinicians intervene earlier in the process of clinical deterioration than they might otherwise, and may prevent asthma exacerbation, emergency department visits, and emergency hospitalization.

Study Population

The study was conducted with the patient-users of the two automated home-based health care systems. Patients for the TLC-DietAid evaluation were under treatment for dyslipidemia at two university hospital primary care clinics and two primary care practices in the community. Patients for the HAT evaluation were being treated at an asthma clinic at a university hospital. The eligibility criteria included (1) age 20 and older; (2) physician diagnosis of dyslipidemia or asthma; (3) understanding of spoken English. The names and telephone numbers of patients were provided to the investigators by their clinicians. The Institutional Review Board approved both evaluation studies and all patients signed an informed consent document.

Direct Observation and In-Depth Interviews

To perform the evaluation studies, we used direct observation and semistructured, in-depth interviews of selected patient-users. We evaluated a total of

13 patients (8 for DietAid and 5 for HAT). We enrolled study subjects for the evaluation of each system until nothing substantially new was being discovered from interviews of the most recent cases. This procedure for setting the sample size for qualitative research studies follows the principles of "information/theoretical saturation."[30,31] Subsequently, the patients used their assigned system at home for 2 weeks at the end of which another in-depth interview was conducted over the telephone.

For each subject evaluated, we first demonstrated how the system worked, trained the subject to use it, observed the person using it, and corrected any mistakes. The in-depth interviews occurred at the end of the session. Because of basic differences in the user interface and content of the two systems, we selected different settings to conduct the training, observation, and the initial interview. For TLC-DietAid, these were done at the study's research offices; for the HAT system, they were conducted in participants' homes.

The training for TLC-DietAid took 10–15 min, and utilized an automated TLC instruction conversation called TLC Training Module. This TLC module explains how TLC works and how the user should interact with the system, and is customized for training individuals to use a particular TLC system. The TLC-DietAid Training Module thus used content from TLC-DietAid in the training. During the training, we used a speakerphone to listen to the interaction between TLC-DietAid and the patient. We were, therefore, able to respond when problems arose that could not be handled by the TLC-DietAid Training Module.

For the HAT system, training took 30–40 min. Patients were instructed on how to operate the PDA and on how to perform a spirometry test in their homes.

Following the training, we gave the subjects a "Users' Guide." The Guide contained information on how to use the system, answers to common questions, and explanations on how to solve problems. Participants were given a toll-free telephone number (Helpline) that they could call if they experienced any difficulty using the system.

Once the demonstration was completed and all questions answered, participants used the system on their own while an investigator observed them. The investigator noted the time the participants spent in performing each of the tasks required of system use, as well as instances of hesitation, indecision, or specific problems encountered. Participants were given ample opportunity to express their opinions about the system and were able to end the interview when they wished. None of the participants, however, ended the session prematurely.

Once the initial interview was completed, participants were asked to use the system at home every day for 2 weeks. Patients were told that they would have an in-depth follow-up telephone interview at the end of the 2-week period. The duration of the follow-up interview was between 15 and 45 min depending upon how much the participants wished to communicate.

During interviews, we followed an "interview guide approach" in which a set of issues to be explored during the interviews with study participants are defined in advance in an interview guide. The interview guide served as a question checklist for the interviewer to make sure that all relevant topics are covered. Using established procedures of qualitative interviewing[30,31] immediately after each conversation the

collected data were analyzed and the results were used to generate interview questions for the next study subject. Thus, the interview protocol evolved during the course of the interviews. This method was also used during the follow-up telephone interviews. The last step in the evaluation process was the reexamination and integration of all the users' responses into a comprehensive set of observations, conclusions, and recommendations for modifying each system on the basis of the evaluation.

RESULTS

The analysis of the in-depth interviews uncovered two major categories of patient concerns: (a) medium-specific issues, i.e., those that dealt with the participants' reactions to the structural aspects of each system's interface with the user such as ease of use, understanding of instructions, etc., and (b) content-specific issues, i.e., those that dealt with the participants' reactions to the messages delivered by the systems. For both systems, medium-specific issues were detected mostly during the initial training, observation, and interview session. Most content-specific issues, however, were brought up during the follow-up telephone interviews. It seems that the 2-week period of time devoted to the utilization of both systems by the participants provided them with ample time to reflect on the specific features of the systems.

Medium-Specific Issues for TLC-DietAid

We found out that the majority of the TLC-DietAid participants experienced the same problems in using the system which supported a conclusion that there were inherent problems with the design and operation of the system, and not idiosyncratic issues for particular users. These problems were of two different types: (1) issues related to users' cognitive capabilities, and (2) issues related to the system's design. The first type of issue became apparent when all participants except one had problems following a number of the instructions for using TLC-DietAid. Even though both the TLC Training Module and the Users' Guide provided instructions to patients on how to use the system, many patients either misinterpreted the instructions or found it difficult to remember them. For example, in response to TLC-DietAid's questions, users were asked to press the number "1" key on their phone to signify "yes" and the number "2" key to signify "no." This worked well as throughout the conversation the two keys (1 and 2) were designed to be used to express the same answer to a question and thus were easily remembered. However, in a few sections of the conversation, the system instructed the participants to press other numbers such as "0" and "99," thus adding a few more options to the multiple choice responses. Subsequently, during our observations, we realized that the majority of the participants had difficulty remembering the key that signified a particular option. As a result, we modified TLC-DietAid so that the system repeated the options a few times particularly in the beginning of the conversation when these concepts had been newly introduced. We also realized that we needed to emphasize the system's "Help" functions.

We identified a similar problem that affected the patients' ability to use TLC-DietAid in a section of the intervention that quizzed them about their knowledge of

nutrition principles. In this section, TLC-DietAid asked them to enter their responses to the questions as either "true" or "false," but failed to explain that "1" stood for "true" and "2" stood for "false" as this was assumed by the system's designers. All participants except one hesitated to enter a response to these questions as they were unsure how to proceed. As a result, we clarified the instructions given to users in that section of the TLD-DietAid.

We also learned that the system neglected to point out some helpful features to the participants. TLC-DietAid contained two features intended to help users respond to questions. They were (1) repetition of a question if a user failed to enter a response after 5 s, (2) pressing the pound key (#) if the user wished to return to the previous question after the person had entered an answer (to possibly change the answer). These two features were described in the Users' Guide and were reviewed by the Training Module at the time of their training. However, TLC-DietAid itself did not remind the users of these features.

These and other issues were identified and discussed mostly during the initial training, observation, and interview session. During this session, one of the participants discovered a software error in the system, something that the extensive software testing had failed to uncover.

Medium-Specific Issues for the HAT System

Medium-specific issues were brought to our attention during both the initial and follow-up interviews. One of the most important of these was the length of the training session. We learned that each individual participant required a different approach during this session. While younger participants could very well tolerate a 1-h session, the older participants stated that the duration was too long and thus tiresome. They felt that two shorter visits would have been less burdensome for them. Other individuals, particularly those who had never worked with computers, needed and wanted more assistance during the training session on how to use the system. For example, one patient needed help to learn how to push the buttons on the PDA and another wanted more assistance with using the "Enter" key. As a result, we realized that the Users' Guide was too brief, and that we needed to provide a more detailed description on how to use the HAT System in it. We also created an on-line "Help" for the HAT patients that described how the system should be used and how to carry out the spirometry testing. The on-line "Help" also contained responses to common questions. In contrast to our evaluation of TLC-DietAid, we did not uncover any HAT system user interface design problems.

Content-Specific Issues for TLC-DietAid

The participants showed different reactions to the system's dietary behavior change content on the basis of individual preferences. Overall, most participants reacted very positively to the system. During the training and first interview session, most indicated that they "learned a lot." Similarly, during the follow-up interview, after all participants had used the system at home on a daily basis for 2 weeks, all said that the information TLC-DietAid provided was useful to them and that they

were thinking of changing some of their dietary habits on the basis of the system's assessment of their diet and the advice it gave them on how to change their habits. In fact, one woman said that she had begun eating more "fish." In some cases, however, the users suggested that we should add to our list of food items, for example adding "grape nuts" to the list of cereals. Another useful suggestion was to add "healthy options" to choose from in the section on "Eating in the Fast-Food Restaurants" such as choosing BK Broiler (sauce taken out!) at Burger King. On the basis of the reactions of the participants, we added "stir fry" and "grill" options to the section on cooking methods.

Other participants expressed concerns that advice offered to them by the system was inappropriate for them. For example, one participant objected to having orange juice or grapefruit juice among recommended fruit juices since the "acidity" in these juices caused the person heartburn. Another participant suggested that TLC-DietAid should not recommend the consumption of fruits and vegetables that were out of season and thus difficult to find or costly. One user made an observation about the "dining-out" section in which participants are encouraged to choose healthier menu items at restaurants. He pointed out: "most people go out to eat it all" and therefore he did not like the recommendation that suggested modification of his food selection. Another individual indicated that we should not include fish among our food items for him since he did not eat fish at all. Such patient-specific comments led us to modify some parts of the system to better address individual preferences and opinions.

Content-Specific Issues for the HAT System

The participants' critiques of the content of the HAT System were also varied and particular to the specific user as they were for TLC-DietAid. The overall response was extremely positive. All participants indicated that having the system at home made them more aware of their asthma and helped them remember to take their medications on time. One participant said that since she began using the system, the frequency of her asthma attacks declined markedly. A number of issues, however, were raised in relation to the content of the system. Some pointed out that we should use lay terms when asking the asthma status questions on the PDA since it was difficult for them to understand a number of jargons we had used. For example, some of the patients did not understand the words "bronchodialator" and "pulmonology" which prompted us to modify the language to "quick relief inhaler" and "lung disease." One person asked us the meaning of the sentence "the data is now complete." Other "interface" issues were identified by some participants. One patient pointed out that our introductory logo on the PDA screen was too large and wordy and thus distracting. This was subsequently modified. Overall, HAT users had fewer problems than did TLD-DietAid users.

DISCUSSION

This study explored the value of direct observation of users and semistructured in-depth interviews in the usability testing of two automated interactive home-based

systems for disease management in patients with hypercholesterolemia and asthma: TLC-DietAid and HAT. These two systems have similar functional goals to monitor and improve the self-care of patients, but differ in terms of content, computer technology, and user interface.

Our previous experience in using formalized structured questionnaires for evaluation[25,32] had shown that while structured questionnaires are effective in assessing patient attitudes towards predefined issues, they lacked the ability to identify remediable system deficiencies and they did not address patient-specific concerns. Our methodology reflected the fact that the evaluation of medical information systems particularly those utilized by patients cannot be separated from the context of use. As social systems, these systems should be evaluated with a focus on the interaction between the technology and contextual aspects of utilization.[33,34]

Scholarly debate on the appropriate methodology for the evaluation of medical information systems has been underway for more than two decades.[35,36] However, qualitative approaches for evaluation of these systems have gained more recognition only during the past 10 years.[20,34] Among various qualitative approaches to evaluation, ethnography is considered to be the most appropriate by some IS scholars.[21,23,37] However, an ethnographic evaluation of medical information system, though ideally a valuable methodology, is generally neither practical nor affordable as it may be too intrusive in the user's life and expensive in terms of the researcher's time. We adopted an alternative approach, i.e., participant observation and two interviews which offered the advantages of ethnography, i.e., it allowed exploration of the users' experience over time (we observed the users and conducted two in-depth interviews), without its constraints (cost, intrusiveness). It also facilitated the observation of temporally changing behavioral phenomena such as change of food-purchasing habits and adherence to asthma medication regimen.

The results of these two pilot studies demonstrate that observation and semistructured in-depth interviews are valuable tools in determining the usability and acceptability with automated interactive systems in health care.[38] These methods are useful in identifying specific, remediable problems that interfere with the ease of use, acceptability and perceived utility of these systems as well as the errors and correctable deficiencies in both systems. On the basis of our results, it would be reasonable to conclude that all automated systems like the ones evaluated in our studies should be pretested *before* they are released for general use and before they are evaluated in expensive randomized clinical trials.

Furthermore, for the two pilot projects under study here, their utility went beyond generating usability data. We learned that patients do not receive health information in a neutral manner. Indeed, in both pilot studies, patients tended to assess and interpret the information provided to them in a way that best fit their lives and their particular situations. We also observed suggestive evidence that use of the two systems, for as brief a period as 2 weeks, had positive effects on the health behavior of some of the users. As described above, one HAT patient, who liked the HAT system's cues promoting medication adherence, said that by remembering to take her medications she was gaining better control over her asthma. This patient also said that the HAT system had provided her with a new awareness about her condition. Similarly, a patient who participated in the DietAid evaluation pilot study indicated that after

being on the system for 2 weeks, she was more mindful of healthy choices when she went food shopping and that she was eating more "fish" as TLC had suggested. These and other indicators of likely intervention effect supported our decision for both HAT and TLC-DietAid, to evaluate their efficacy in randomized clinical trials.

REFERENCES

1. Slack, V., *How Computing Empowers Doctors and Patients for Better Health Care*, Jossey-Bass, San Francisco, 1997.
2. Eng, T. R., Maxfield, A., Patrick, K., Deering, M. J., Ratzan, S. C., and Gustafson, D. H., Access to health information and support: A public highway or a private road? *JAMA* 280:1371–1375, 1997.
3. Eng, T. R., and Gustafson, D. H. (eds.). *Wired for Health and Well-Being: The Emergence of Interactive Health Communication*, Science Panel on Interactive Communication and Health, U.S. Department of Health and Human Services, Washington, DC, 1999.
4. Friedman, R. H., Stollerman, J., Mahoney, D. M., and Rozenblyum, L., The virtual visit: Using telecommunications technology to take care of patients. *JAMIA* 4:413–425, 1997.
5. Brennan, P. F., Moore, S. M., and Smyth, K. A., The effects of special computer network on caregivers of persons with Alzheimer's disease. *Nurs. Res.* 44:166–508, 1995.
6. Friedman, R. H., Kazis, L. E., Jette, A., Smith, M. B., Stollerman, J., Torgerson, J., and Carey, K. B., A telecommunication system for monitoring and counseling patients with hypertension: Impact on medication adherence and blood pressure control, *Am. J. Hypertens*. 9:285–292, 1996.
7. Alemi, F., Stephens, R. C., Muise, K., Dyches, H., Mosavel, M., and Butts, J., Educating patients at home: Community health rap. *Med. Care* 34:S21–S31, 1996.
8. Friedman, R. H., Automated telephone conversations to assess health behavior and deliver behavioral interventions. *J. Med. Syst*. 22:95–102, 1998.
9. Polson, P. G., Lewis, C., Rieman, J., and Wharton, C., Cognitive walkthrough: A method for theory-based evaluation of user interfaces. *Int. J. Man–Machine Stud*. 36:741–773, 1992.
10. Novick, D. G., Using the cognitive walkthrough for operating procedures. *Interactions* 6(3):31–37, 1999.
11. Jeffries, R., Miller, J. R., Wharton, C., and Uyeda, K. M., User interface evaluation in the real world: A comparison of four techniques. *Proc. 1991 CHI Conf.* New Orleans, ACM, LA, pp. 119–124, 1991.
12. Nielsen, J., and Molich, R., Heuristic evaluation of user interfaces. *Proc. CHI'90 Conf. Hum. Factors Comput. Syst.* New York, *ACM*. pp. 249–256, 1990.
13. Nielsen, J., *Usability Engineering*, Academic Press, Cambridge, MA, 1992.
14. Rowley, D. E., and Rhodes, D. G., The cognitive jog through: A fast-paced user interface evaluation procedure. *Proc. CHI'92 Conf. Hum. Factors Comput. Syst.* Monterey, *ACM*, CA, pp. 389–395, 1992.
15. Nissen, H. S., Acquiring knowledge of information systems—research in a methodological quagmire. In Mumford, E., *et al.* (eds.), *Research Methods in Information Systems*, Elsevier, North Holland, The Netherlands, pp. 39–51, 1985.
16. Galliers, R. D., and Land, F. F., Choosing appropriate information systems research methodologies, *Assoc. Comput. Machinery* 30:900–902, 1987.
17. Vitalari, N., The need for longitudinal designs in the study of computing environments. In Mumford, E., *et al.* (eds.),*Research Methods in Information Systems*. Elsevier, North-Holland, The Netherlands, pp. 243–265, 1985.
18. Kautz, K., User participation and participatory design: Topics in computing education. *Human–Comput. Interact.* 11:268–285, 1996.
19. Gulliksen, J., Lantz, A., and Voivie, I., User centered design—Problems and possibilities: A summary of the 1998 PDC and CSCW workshop. *SIGCHI Bull*. 31(2):25–35, 1999.
20. Wixen, D., and Ramey, J., *Field Methods Casebook for Software Design*, Wiley, New York, 1996.
21. Lewis, S., Mateas, M., Palmiter, S., and Lynch, G., Ethnographic data for product development: A collaborative process. *Interactions* 3(6):52–69, 1996.
22. Beyer, H., and Holtzblatt, K., Contextual design. *Interactions* 6(1):32–42, 1999.
23. Prasad, P., Systems of meaning: Ethnography as a methodology for the study of information technologies. In Lee, A. S., Leibenau, J., DeGross, J. I. (eds.). *Information Systems and Qualitative Research*, Chapman & Hall, London, pp. 101–118, 1997.
24. Forsythe, D., Using ethnography in the design of an explanation system, *Expert Syst. Appl.* 8:403–417, 1995.

25. Finkelstein, J., Cabrera, M. R., and Hripcsak, G., Internet-based home asthma telemonitoring: Can patients handle the technology? *Chest* 117:148–155, 2000.
26. Pearson, T. A., Brown, W. V., and Donato, K., *et al.*, Lipids, Workshop IV. *Circulation* 88:1397–1401, 1993.
27. Bandura, A., *Social Learning Theory*, Prentice-Hall, Englewood Cliffs, NJ, 1977.
28. Expert Panel Report 2. *Guidelines for the Diagnosis and Management of Asthma*. National Institutes of Health, National Heart, Lung and Blood Institute, NIH Publication No. 97–4051, April 1997.
29. Finkelstein, J., and Cabrera, M. R., Telematic system for monitoring of asthma severity in patients' homes, *Med. Inform.* 9(pt 1):272–276, 1997.
30. Strauss, A., and Corbin, J., *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*, Wiley, New York, 1984.
31. Strauss, A., *Qualitative Analysis for Social Scientists*, Cambridge University Press, New York, 1987.
32. Glanz, K., Shigaki, D., Farzanfar, R., Pinto, B., Kaplan, B., and Friedman, R., Participants' reactions to a computerized telephone-based system for nutrition and exercise counseling. *Patient Educ. Couns*. 49:157–163, 2003.
33. Symons, V., A review of information systems evaluation: Content, context and process. *Eur. J. Inf. Syst.* 1:205–212, 1991.
34. Trauth, E., *Qualitative Research in IS: Issues and Trends*, Idea Group Publishing, New York, 2001.
35. Mumford, E., *et al.* (eds.), *Research Methods in Information Systems*, Elsevier, North-Holland, The Netherlands, 1985.
36. Heathfield, H., Current evaluations of information technology in health care are often inadequate. *BMC* 313:1008, 1996.
37. Myers, M., Investigating information systems with ethnographic research. *Commun. AIS* 2:2–19, 1999.
38. Heathfield, H., Evaluating information technology in health care. *BMJ* 316:1959–1961, 1998.

MEDINFO 2004
M. Fieschi et al. (Eds)
Amsterdam: IOS Press

Computer-Based and Live Interviews on Problem Drinking: Users' Attitudes

Ramesh Farzanfar[a], Jeffrey Migneault[a], Amy Rubin[b], Robert Friedman[a], Lisa Marks[a]

[a] *Medical Information Systems Unit, Boston University & Boston Medical Center, Boston, MA*
[b] *Department of Psychiatry, Massachusetts General Hospital, Boston, MA*

Abstract

Some studies of computerized interviews particularly those that deal with personally sensitive topics demonstrate that people have a preference for automated interviews versus live interviews. To explore this phenomenon, we administered four open-ended questions after participants were screened for problem drinking by both an automated and a human telephone interviewer. Both interviews administered AUDIT (Alcohol Use Disorders Identification Test) for assessing problem drinking. Individuals were recruited into the study who responded to ads in daily papers. Sixty-two percent of the participants preferred the human interviewer and only 3% among these expressed a concern about confidentiality of the interview. Among the 22% who preferred the automated interview, 32% indicated confidentiality as a reason for their preference.

Keywords:

Automated alcohol screening, Human-Computer Interaction, Technology assessment.

Introduction

The utilization of computers to interview patients for symptoms assessment, screening, education, counseling, etc., is becoming more common in medical practice[1,2]. One use of computers is to help patients (and their health care providers) determine whether they are involved in risky health behavior [3,4]. Various methods of assessing risky health behavior including computer-based interviewing have been proposed and evaluated [5]. In fact, some evaluations of computer-based interviewing versus professional interviewers and other methods of data collection for assessment of risky health behavior have demonstrated that this method is effective and may even elicit more truthful responses from participants [6,7,8]. This paper presents the results of analysis of open-ended additional questions administered to individuals who were interviewed separately by a telephone-based human interviewer and a telephone-based computer-controlled automated interview on the subject of problem drinking (people who drink above recommended guidelines or who have alcohol-related problems that do not meet criteria for a diagnosis of alcohol abuse or dependence). In both types of interviews, the AUDIT, a validated screening test for problem drinking was administered. The two interviews were identical in all respects. The open-ended questions were asked by a trained human interviewer via the telephone at the end of the second AUDIT interview. The open-ended additional questions evaluated the participants' overall opinions about the two interviews, their assessment of both methods of interviewing, their preference for one or another method and why, and the truthfulness of their reporting of problem drinking behavior.

Telephone-Linked Care (TLC) is an interactive voice response (IVR) and speech recognition (SR) system that is usually utilized for disease prevention and health promotion and for chronic disease management [9-13]. While most TLC applications are designed to function as a health care intervention, TLC-AUDIT uses a standardized, valid and reliable instrument to screen users for problem drinking [14]. The AUDIT is a 10-item screening tool developed and used by World Health Organization [15] in multinational trials of brief alcohol interventions. The AUDIT produces a total score from 0-40 and has been shown to be generalizable across cultural characteristics [16]. A score of 8 or more indicates problem drinking. In our study, all participants who scored higher than 8 on either the TLC computer or human AUDIT interview were provided with information and advice and were referred to a hotline telephone number for information about problem drinking and its treatment. This information was provided by both the automated interview and the live interviewer.

Methods: Evaluation study design

TLC-AUDIT was evaluated for validity in two separate interviews (a week apart) by comparing users responses to TLC-AUDIT and to a human AUDIT interviewer (N=100). The order of presentation of the two different versions (human or computer) was random. At the end of the second AUDIT interview, we evaluated users' attitudes towards the two interviews by administering four open-ended questions. The open-ended questions were administered by the live interviewer. The questions addressed the following topics: 1) participants' thoughts or feelings about the two AUDIT interviews in their own words, 2) the accuracy of the participants' responses, 3) whether the "right" questions were asked, 4) participants' preference for the computer or the live interviewer.

Participants' responses were transcribed and a subsequent content analysis of the responses to each question was carried out. The analysis had both quantitative and qualitative dimensions. A systematic coding of the text led to the generation of thematic categories and sub-categories. Continued analysis entailed fre-

quent revisions and reconsiderations of categories and sub-categories. Subsequently, the frequencies of coded concepts were counted [17]. The analysis focused on identifying both recurrent and unique experiences as well as some of the dimensions and context of those experiences. At the same time, each category was compared and contrasted with clusters of other categories in the text [18].

Results

Question 1- Overall thoughts or comments

Participants were eager to share their thoughts as over 90% did in fact express their opinions about the interviews. Their comments addressed either both interviews or the automated interview: A) Some comments were expressed in general terms and thus directed towards both interviews without explicitly addressing one particular interview (live or automated). Twenty five percent of these comments were positive. Some positive comments revealed an interesting dimension that was expressed as a perception of an increased "awareness." A response such as "the questions made me think" was expressed by several people. Another dimension was captured in the statements that pointed out the informational value of the interviews. "Informative" was thus repeated by several other individuals. Four individuals expressed negative views about the two interviews. An example is the person who said that the whole interview process was "too long" adding "one needs only an IQ of 60 to answer these questions!" There were also responses directed toward both interviews that were not explicitly positive but did not have an overtly negative tone either (neutral comments=50%). Within this category a theme emerged that essentially addressed the design of the structured interview questionnaire and the multiple choice responses provided as answers. Over 10% of the participants in this category objected to the close ended multiple choice questions. One comment highlights this objection: "it was difficult to choose one of the allowable responses." This sentiment, directed either at the structured close-ended questions or the multiple choice answers, was repeated by several respondents: "some of my answers were incorrect. I didn't mean the [every month] answer," "the choices were not specific enough." One notable comment went so far as saying "response options were better with the interviewer!" It seems that for this person interaction with the live interviewer had a mitigating effect on the perceived complexity and perhaps difficulty of answering the identical close-ended questions administered by TLC-AUDI. B) Comments directed specifically at the automated interview. Ten percent of the participants expressed positive responses towards TLC-AUDIT interview. For example, "the computer was very good," "the computer was well programmed...because you could interrupt it to give answers," etc. It is interesting, however, that later during the interview, nearly half of those respondents who had initially expressed positive opinions about the automated interview added that they preferred to be interviewed by a person. Only two individuals had negative comments about the automated interview.

Question 2- Accuracy

Almost all participants responded to this question. The category of "accuracy" was divided into three sub-categories: 1) completely accurate, 2) mostly accurate, and 3) possible (some) inaccuracies. All those who used the words "very," "extremely," "100%," to describe the accuracy of their response and/or used the word "accurate" without qualification were placed under "completely accurate" sub-category (50%). Participants who qualified their description of accuracy with such words as "pretty" or "fairly" accurate were placed under 'mostly accurate" subcategory (23%). The subcategory "possible inaccuracies" included those responses that admitted to some inaccuracies or said that they were "80%," "85%," or "90%" accurate (19%).Those who said that there may be some inaccuracies in their response mostly attributed the reasons to problems with the computer or the limitations in the close-ended, multiple choice questionnaires. One participant mentioned that the computer "was on autopilot," and that "the timing on the computer made me a little antsy because I felt I didn't have time to think." Or, a 65 year old female who had problems with the speech recognition and thus used the keypad to enter her answers, felt that her responses were "probably not accurate."

Question 3- Which, computer or live interviewer, asked the right questions?

In response to this question, 57% of the participants replied that the questions were the same in both interviews [the "correct" response], while 27% said that the live interviewer asked the "right" question. Only three persons (3%) said that "the computer asked the right questions" or that they "liked the questions by the computer." The rest of the responses were too disparate to generate a conceptual category and were grouped under "can't say/no response," and "Other" categories with such examples as "both did OK." Within this group two persons actually said that they "liked the questions by the computer." Furthermore, Twenty-eight percent of the respondents indicated that additional questions should have been asked. The most frequently suggested question was "when do you drink?" By "when," these participants meant either: 1) the timing of their drinking, e.g., weekends, or 2) circumstances of their drinking, e.g., when depressed. One person felt that "both interviews were problematic" because she did most of her drinking during the weekends. This respondent whose AUDIT scores were 18 and 17 respectively, felt that if the pattern of her drinking habit was taken into account, her scores would have been lower. This sentiment was echoed by 10 other participants. One participant, a 69 year old male, said "I only drink on weekends when I'm out all day. I never drink during the week and never drink much at home." This participant's AUDIT scores were 26 and 22 respectively, and he happened to mention that his doctor had told him "to cut back." Also, a 22 year old male mentioned that the interview didn't "differentiate between long term and short term drinking" since "college students binge [drink] over the summer." Without exception, all participants who made such temporally related suggestions had AUDIT scores above 8. Some respondents who suggested adding "when do you drink," intended the question to address the reasons why people drink. An example is the following comment: "people drink for different reasons and what con-

dition they're in and if they're depressed. There should be questions to assess this."

Question 4- Interviewer Preference

The responses to this question were divided into two broad categories: "favorable to the live interview" and "favorable to the automated interview." Sixty two percent of respondents preferred the live interviewer, while 22% said that they preferred the automated interview and 15% believed there was no difference between the two. One person gave no explicit answer to this question. The subcategories within the category of "favorable to the live interview" included the following: "personal/ easy to talk to/more comfortable," (37%), "aversion to computers," (19%), "gives feedback and clarifications," (11%), "no reason provided," (11%), "confidential/anonymous" (3%), and "other," (15%). Two individuals said they preferred the live interviewer but also added that it "depends on interviewer." Examples of several notable responses in these subcategories include: "I need interaction with people with heart," "I like the personal touch," "A person gives me a chance to say more," etc. Among the people who preferred a live interviewer, 31% thought that the live interviewer asked them the "right" questions. Also, 16% indicated that their interview with the computer was problematic such as "A person understands me better," "computer is confusing. I couldn't find the right answer." As suggested above, some respondents framed their preference for the live interviewer in statements that conveyed their disapproval of the automated interview. Examples include: "I prefer the person because computer is fake; it is a machine," "Computer only allows yes or no answers," "With computers, you can't ask a question," "Talking to a computer is like talking to a wall."

Similarly, it seems that for some participants expression of a favorable response toward the live interviewer was indicative of perceived problems with the computer. In fact 14% of respondents who preferred the live interviewer reported having some problems with the automated interview. A 64 year old male said that "with the computer I had to think more about the computer than the ideas about drinking." Similarly, a 20 year old female said that she was worried that thinking about her answers "might cause a problem and then I might skip a question." Likewise, a 29 year old male said that he "got mixed up a little bit. It doesn't stop, just keeps going; if you miss, the chance is gone!" Finally, two people who preferred the live interviewer thought that only she had given them the hot line number! Both persons had their second interview with the live interviewer. Therefore, it is possible that they were better able to recall their conversation with the live interviewer even though the computer also had provided them with the hot line number. Thirteen individuals complained that the computer could not understand them very well. Five among these said that they had to speak too loud so that the computer could understand them. And, three people specifically expressed frustration with the automated interview. Interestingly, among those who preferred the automated interview (22% of the study population) only three individuals said that they had a problem with the computer. The reasons provided by those who preferred the automated interview included "more privacy/confidential" (32%) and "non-judgmental" (18%). The most compelling statement came from one participant who said "I've been fighting with therapists all my life. I've been getting up and leaving. I didn't feel I had to with this [the computer]." An estimated 23% of those who preferred the automated interview actually said that the live interviewer asked the "right" questions. At the same time, 23% said that an automated interview should be followed by an interview with a live person. An interesting dimension of the responses in this category was the ambivalence expressed by a few. One respondent said that "I prefer to be interviewed by the computer but I think a person would be more accurate and have more feeling." Another person first expressed her preference for the "human touch" and then said that she prefers the computer "because it doesn't talk back!"

There were also those participants who said they had no preference between the computer and the live interviewer or felt that the experience was the same (15%). Responses in this category ran the gamut of "no preference" to more detailed comments. Most of the respondents who articulated a reason tried to explain why both interviewing methods were appropriate such as the following: "A person explains more in detail, but computer offers more privacy." "Over the phone questions can be asked by anyone about anything because there is no one looking at me and judging me." Two persons, in fact, indicated that after having experienced both the automated and the live interviews, they realized that in the future they would be comfortable with both. Finally, two individuals in this category thought the interviewer asked the "right" questions as well as "more" questions.

Discussion

Our results are at variance with those presented by some investigators about people's preference for computers over live interviewers when responding to personal and sensitive questions [4,6]. Several factors might have contributed to the findings of this study: 1) the live interviewer's communication skills and sociability might have been a factor that contributed to the formation of a "favorable" opinion towards the live interviewer. 2) Social desirability is another factor to consider as participants might have wanted to please the interviewer by saying that they preferred her over the computer. In fact, one participant after stating that he had no comments, immediately added "the interviewer was nice!" 3) Possible problems with the speech recognition may be another reason for the participants' preference for the live interviewer. Among the respondents who preferred the live interviewer, 16% (10) had problems with the automated interview. 4) Some studies that have provided evidence of users' preference for computer-based interviews, have made comparisons between automated interviews and traditional face-to-face interviews [19,20]. The fact that in our study both AUDIT interviews were conducted over the phone seems to have diminished the awkwardness or discomfort of divulging personal information. As one participant said, "as long as the interview is over the phone, it could be a person or a computer – no difference to me." 6) We recruited our study participants by posting advertisements on daily papers. The recruitment method was thus a self-referral one as participants were encouraged to call to see "whether they

are problem drinkers." The very nature of such an act, i.e., to make an attempt to find out whether one in involved in risky health behavior, is charged with emotions. For example, a 66 years old female participant who, after the AUDIT interview, learned that she is not a problem drinker said: "I'm glad I'm not at high risk. I feel relieved." Indeed, a number of participants seemed to be seeking help or thought that the study would be providing some help. These individuals sought to communicate with a live person, someone to offer help, perhaps some counseling, a sympathetic ear, etc. One individual actually asked the interviewer "offer treatment?" Or, another participant who admitted he had been to "detox," said that he preferred a person because he "could tell more." Another participant said: "someone might do the computer interview but not follow through and get further help."

This evaluation study explored the opinions of users of computer-based and human interviews on problem drinking. Our results differ from some studies that report users' preference for automated interviews versus human interviews and raise the question that users' preferences for an automated interview versus a live one is a complex issue.

References

[1] Slack W. Cyber Medicine: *How computing empowers doctors and patients for better health care.* San Francisco: Jossey-Bass Publishers, 1997.

[2] Doolan DF, Bates DW, Brent CJ. The use of computers for clinical care: a case series of advanced US sites. *J Am Med Informatics Association* 2003; 10:94-107.

[3] Gerbert B, Bronstone A, Pantilat S, McPhee S, Allerton M. When asked, Patients tell: disclosure of sensitive health-risk behaviors. *Med Care* 1999; 37:104-111.

[4] Tourangeau R, Smith TW. Asking sensitive questions: the impact of data collection mode, question format and question context. *Pub Opin Q* 1996; 60:275-304.

[5] Richter L, Johnson PB. Current methods assessing substance use: a review of strength, problems and developments. *J Drug Issues* 2001; 31:809-832.

[6] Locke SD, Gilbert BO. Method of psychological assessment, self-disclosure, and experimental differences: a study of computer, questionnaire and interview assessment formats. *J Soc Behavior and Personality* 1995; 10:255-263.

[7] Yates B, Wagner J, Supernane L. Recall of health-risky behaviors for the prior 2 or 4 weeks via computerized versus printed questionnaire. *Comput Hum Behavior* 1997; 13:83-110.

[8] Lessler JT, Caspar RA, Penne MA. Developing computer-assisted interviewing (CAI) for the National House Survey on Drug Abuse. *J Drug Issues* 2000; 30:9-33.

[9] Friedman RH, Kazis LE, Jette A et al.A telecommunication system for monitoring and counseling patients with hypertension: impact on medication adherence and blood pressure control. *Am J Hypertens* 1996; 9:285-292.

[10]Jarvis K, Friedman R, Heeren T, Cullinane P. Older women and physical activity: using the telephone to walk. *Womens Health Issues* 1997; 7:24-29.

[11]Friedman RH. Automated telephone conversations to assess health behavior and deliver behavioral interventions. *J Med Syst* 1998; 22:95-102.

[12]Ramelson H, Friedman R, Ockene JK. An automated telephone-based smoking cessation education counseling system. *Patient Educ Couns* 1999; 36:131-143.

[13]Delichatsios H, Friedman RH, Glanz K, Tennstedt S, Smigelski C, Pinto BM, Kelley H, Gillman MW. Randomized trial of a 'talking computer' to improve adults' eating habits. *Am J Health Promot* 2001; 15:215-224.

[14]Higgins-Biddle JC, Babor TF, Mullahy J, Daniels J, McRee B. Alcohol screening and brief intervention: where research meets practice. *Conn Med* 1997; 61:565-575.

[15]Babor TF, de la Fuente JR, Saunders J, Grant M. *AUDIT: The Alcohol Use Disorders Identification Test: guidelines for use in primary health care.* Geneva: Switzerland: World Health Organization, 1992.

[16]Allen JP, Litten RZ, Fertig JB, Babor TF. A review of research on the Alcohol Use Disorders Identification Test (AUDIT). *Alcohol Clin Exp Res* 1997 21:613-619.

[17]Pool Ithiel de Sola, *Trends in Content Analysis.* Urbana, IL: University of Illinois Press, 1959.

[18]Mostyn B. The content analysis of qualitative research data: a dynamic approach. In: Brenner M, Brown J and Cayker D, eds. *The Research Interview.* London: Academic Press, 1985; p. 115-145.

[19] Des Jarlais DC, Paone D, Milliken J, Turner CF, Miller H, Gribble J, Shi Q, Hagan H, Friedman SR. Autio-computer interviewing to measure risk behavior for HIV among injecting drug users: a quasi-randomized trial. *Lancet* 1999; 354:678-679.

[20]Gribble JN, Miller HG, Cooley PC, Catania JA, Pollack L, Turner CF. The impact of T-ACASI interviewing on reported drug use among men who have sex with men. *Subst Use Misuse* 2000; 35:869-890.

Address for correspondance

Ramesh Farzanfar, PhD
Medical Information Systems Unit
Boston Medical Center
720 Harrison Avenue , Suite 1102
email: rfarzanf@bu.edu

TRENDS

Place Of Death: U.S. Trends Since 1980

Fewer Americans died in the hospital in 1998 than in 1980, but some racial disparities raise troubling questions.

by James Flory, Yinong Young-Xu, Ipek Gurol, Norman Levinsky, Arlene Ash, and Ezekiel Emanuel

ABSTRACT: Place of death is one indicator of the state of end-of-life care. We examine trends in national death certificate data on place of death from 1980 to 1998. During these years the percentage of Americans dying as hospital inpatients decreased from approximately 54 percent to 41 percent. About 310,000 fewer people died in the hospital in 1998 than if the proportion of inpatient deaths had not changed since 1980. For certain diseases the change was much greater. In 1980 whites and African Americans died in the hospital in equal proportions, but in 1998 whites died as inpatients less often than African Americans. These racial differences and their implications deserve further study.

DURING THE PAST two decades end-of-life care has become an increasingly high-profile issue.[1] In the 1980s Medicare introduced the hospice benefit, and numerous judicial rulings expanded patients' right to terminate life-sustaining interventions.[2] In the 1990s the United States Supreme Court rendered three decisions on end-of-life care; the Patient Self-Determination Act was enacted; major research projects to identify and overcome barriers to better end-of-life care were conducted; and initiatives were undertaken to expand hospice and educate physicians about end-of-life care.[3] Other events not specifically directed at end-of-life care also might have altered such care. These include the introduction of capitation in Medicare, the expansion of Medicare's postacute care benefit, the overall decline in the use of inpatient care, the expansion of managed care, and the aging of the population.[4]

A key issue in end-of-life care is where people die. The location of death shows where the patient was receiving care at the very end of life and suggests what could be done to improve that care. If most people die in hospitals, then hospital policies and hospital staff training are particularly important. If more people die in nursing homes or skilled nursing facilities, these institutions' ability to deal appropriately with dying becomes more important. If deaths occur in homes, the quality of home care becomes more pressing.

Also, many people in the end-of-life care community look upon an increase in home deaths favorably, partly because of evidence

James Flory (jflory@cc.nih.gov) is a predoctoral fellow in the Department of Clinical Bioethics, Warren G. Magnuson Clinical Center, National Institutes of Health, in Bethesda, Maryland. Yinong Young-Xu is an epidemiologist in the Section of General Internal Medicine, Boston Medical Center, in Boston, Massachusetts. Ipek Gurol is an assistant professor in the Department of Economics, Bogazici University, Turkey. Norman Levinsky, who died in March 2004, was a professor and associate provost of the Department of Medicine, Boston University School of Medicine, where Arlene Ash is a research professor. Ezekiel Emanuel is chair of the NIH Department of Clinical Bioethics.

DOI 10.1377/hlthaff.23.3.194

that a large majority of Americans prefer to die at home.[5] Some of the most prominent efforts to improve end-of-life care, such as hospice, are meant to facilitate death at home or in homelike surroundings.[6] Changes in the place of death could indicate whether these efforts have had any impact. This paper seeks to inform discussion of end-of-life care policy issues by examining trends in place of death in the context of changes in the health system.

Study Methods

The National Vital Statistics System (NVSS) has collected death certificate records from all fifty states since 1980. Death certificates provide information on place of death, as well as cause of death; region/state/county of death; and the age, sex, and race of the deceased. We have analyzed NVSS data for the years 1980–1998.

Prior to 1989, place-of-death category codes used in death certificates were as follows: (1) hospital—inpatient, (2) hospital—outpatient or emergency room, (3) hospital—dead on arrival, (4) hospital—patient status unknown, (5) hospital—patient status not on certificate, (6) other institutions, (7) dead on arrival—hospital name not given, (8) hospital and patient status not stated, and (9) all other reported entries. Because of incomplete coding practices before 1989, a large number of U.S. counties coded most or all of their in-hospital deaths as "hospital—patient status not on certificate," failing to distinguish among inpatient, outpatient, dead on arrival, and emergency room deaths. To create reasonably comparable statistics from 1980 through 1998, we included only data from county-year combinations with fully distinct records of inpatient deaths. In the original data the total number of counties for 1980–1988 was 28,148, covering 18.5 million deaths. We have eliminated all counties with any reported "hospital—patient status not on certificate" deaths, leaving 22,689 counties (81 percent) and 12.7 million deaths (69 percent).

Place-of-death coding has been uniform throughout the United States since 1989. The "hospital—patient status not on certificate" code has been eliminated. In addition to the four well-defined hospital categories (1–4) listed above, nursing home and residential deaths have been identified as such, and all other deaths have been recorded either under a single "other" category or as "place of death unknown." The sample for 1989–1998 includes all of the 22.5 million deaths for the period, adding to a total of 35.2 million observations for 1980–1998. In the results we present here, the "hospital—inpatient" category is used as the basis of analyses.

We studied trends in the proportion of inpatient deaths by age, sex, race, cause of death, and region of death across the time period. Coding for the detailed race of the deceased was uniform in the two subsample periods of 1980–1988 and 1989–1998, so we have grouped the racial information into three main categories: white, black, and other. Although post-1989 data have details on Hispanic origin, the Hispanic population has been included in the "white" category for consistency across years. Causes of death on the certificates are recorded with *International Classification of Diseases*, Eighth Revision (ICD-8, pre-1989) and ICD-9 (Ninth Revision, post-1989) codes. We have distinguished eight leading causes of death: chronic heart disease, acute myocardial infarction (AMI), cancer, chronic obstructive pulmonary disease (COPD), stroke, pneumonia and influenza, diabetes, and chronic liver disease. All other causes have been grouped as "other." Regional analyses compared western, midwestern, southern, and eastern regions of the United States. Counties were also distinguished using the urban/rural continuum code developed by the U.S. Department of Agriculture's Economic Research Service.[7]

The only statistical tool in our analysis is testing for the significance of differences between proportions. Because of the number of observations, all differences in proportions reported in Exhibit 1 are statistically significant at $p < .001$. For finer subgroups reported in the text, the sizes of the samples are still large enough to guarantee that any difference in inpatient death proportions at least as great as 2 percent is statistically significant at $p < .05$.

EXHIBIT 1
Demographic Data And Inpatient Death Rates, United States, 1980 And 1998

		Percent of deaths classified as inpatient	
	Percent of population	**1980**	**1998**
All observations (N = 35.2 million deaths)	100	54	41
Age (years)			
<65	25	52	42
65–74	20	60	47
75–84	29	56	44
85+	26	46	34
Sex			
Male	49	54	42
Female	51	54	40
Race			
White	86	54	40
Black	12	54	48
Cause of death			
Chronic heart disease	23	44	35
Acute myocardial infarction	9	44	40
Cancer	23	70	37
Chronic obstructive pulmonary disease	5	66	47
Stroke	7	59	50
Pneumonia and influenza	4	66	60
Diabetes	3	56	41
Chronic liver disease	1	74	60
Region			
Northeast	20	56	44
Midwest	24	53	39
South	37	54	44
West	19	49	36

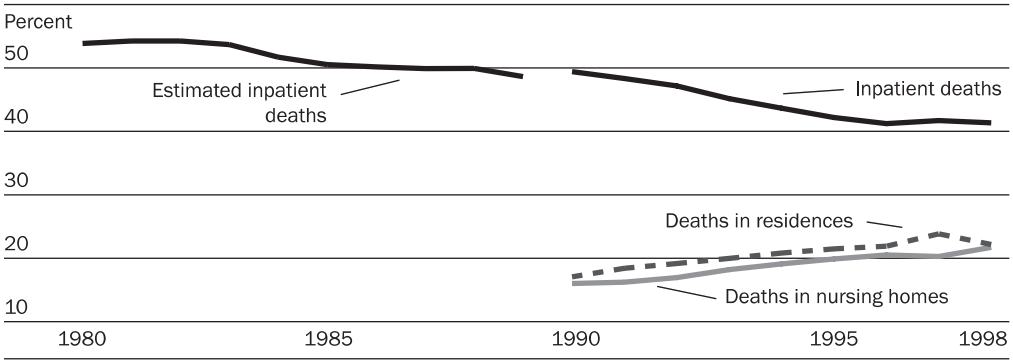
SOURCE: National Vital Statistics System Death Certificate Records.

Study Results

The percentage of persons dying as inpatients held steady from 1980 to 1983. Since 1983 the rate of in-hospital deaths in the United States declined at a fairly constant rate, from a high of about 54 percent to a low of around 41 percent in 1998 (Exhibit 2). During the 1990s the decline was 8 percent—nearly 1 percent a year. Deaths occurring at home and in nursing homes correspondingly increased. Between 1990 and 1998, home deaths rose from 17 percent to 22 percent, and nursing home deaths, from 16 percent to 22 percent. The remaining deaths occurred mainly in outpatient medical facilities (7 percent) and other unspecified locations (4 percent). Patients who were dead on arrival or with status not specified on the death certificate constituted about 3 percent of deaths recorded in the 1990s.

While the proportion of in-hospital deaths fell for each major cause of death, place of death has changed most for cancer (Exhibit 3). Between 1980 and 1998 the rate of inpatient deaths from cancer declined from 70 percent to 37 percent. The percentage of patients with diabetes and COPD who died in the hospital fell by more than 15 percent each (Exhibit 1), as deaths attributable to AMI dropped less

EXHIBIT 2
Decline In Percentage Of Americans Dying As Inpatients, 1980–1998, And Percentage Of U.S. Home And Nursing Home Deaths During The 1990s



SOURCE: National Vital Statistics System Death Certificate Records.

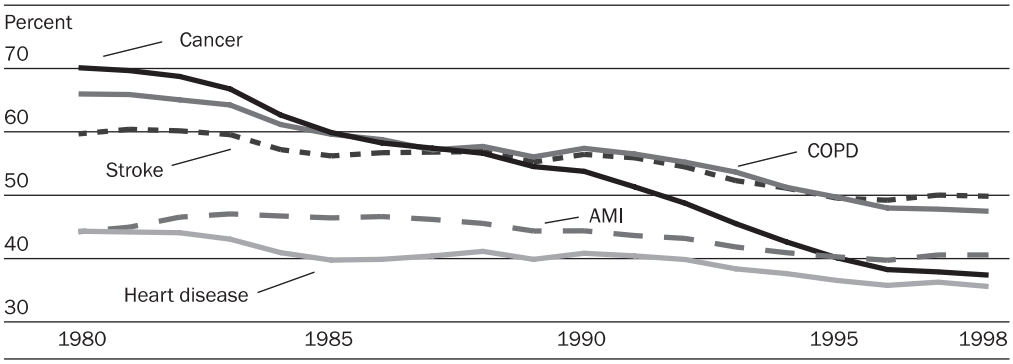
than 4 percent.

Cancer is also unusual in that data from the 1990s show that most of the shift out of the hospital has been toward private residences, which increased by 15 percent from 1989 to 1998 to account for 38 percent of all cancer deaths; in the same time period, cancer deaths in nursing homes rose just 4 percent, to 17 percent. For other diseases, declines in inpatient deaths were evenly divided between increases in home and nursing home deaths.

In 1980 there was almost no difference in rates of inpatient death between whites and African Americans (Exhibit 4). However, by 1998 whites died in the hospital much less frequently (40 percent) than African Americans did (48 percent). The proportion of in-hospital deaths declined for both races for every major cause of death, yet for each cause it declined more for whites.

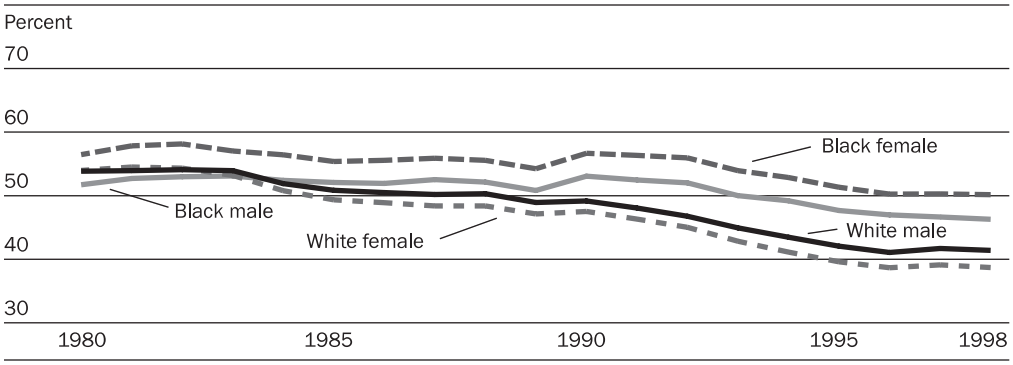
During the 1980s and 1990s the change in in-hospital mortality was similar for both sexes. However, the comparison between sexes differed for African Americans and whites (Exhibit 4). African American and white women had especially different prospects of dying in the hospital. In 1998 white women died in the hospital 39 percent of the time, while African American women did so 50 percent of the time.

EXHIBIT 3
Decline In Percentage Of Americans Dying As Inpatients, By Cause Of Death, 1980–1998



SOURCE: National Vital Statistics System Death Certificate Records.
NOTES: AMI is acute myocardial infarction. COPD is chronic obstructive pulmonary disease.

EXHIBIT 4
Rates Of Inpatient Death By Race And Sex, 1980–1998



SOURCE: National Vital Statistics System Death Certificate Records.

With regard to age, the proportion of in-hospital deaths was highest for deceased people ages 65–74 years, 47 percent in 1998 (Exhibit 1). Nursing home deaths increase greatly with advanced age. In 1998, 24 percent of people ages 75–84, and 45 percent of those older than age 85, died in nursing homes. These relationships between age and place of death did not change much during the study period.

Rates of in-hospital death differed around the country. Inpatient death in 1998 was more likely in the Northeast and South and less likely in the West and Midwest (Exhibit 1). Analysis according to how urban a county was (as classified by the urban/rural continuum code) did not point to any trend variation over time. Inpatient deaths declined at roughly the same rate in metropolitan and nonmetropolitan counties.

Discussion And Policy Implications

During the last two decades of the twentieth century there was a noticeable change in the way Americans died. The hospital ceased to be the setting of death for most Americans. By 1998 more Americans died at home or in a nursing home (45 percent) than died as hospital inpatients (41 percent). Since more than 2.4 million Americans die each year, the substantial shift in the place of death since 1980 means that in 1998 approximately 310,000 people died outside the hospital who would have died in the hospital if the distribution had been the same as it was two decades earlier. Importantly, this shift in place of death occurred relatively steadily during the 1980s and 1990s. This suggests that no single event during that period triggered the change. Reductions in in-hospital deaths have been evolutionary, not revolutionary.

This shift out of the hospital has implications for ensuring that dying patients receive good end-of-life care. Moving outside the hospital setting does not ensure a good death. It could represent movement to a skilled nursing facility, where the environment and the issues for quality of dying might be similar to those of a hospital. Hospices, homes, and traditional nursing homes have their own potential problems. Thus, one should conclude not that end-of-life care has necessarily improved but rather that quality of care in places other than hospitals is becoming more important.

Death certificate records indicate that approximately 500,000 people die annually in nursing homes. Nursing home staff need adequate resources and training to fill their increasingly important role in end-of-life care. In addition, as hospices and home care agencies rapidly expand to meet the needs of the growing numbers of Americans dying in their own homes, quality of care must not be allowed to erode.[8]

■ **Racial differences.** The growing gap in in-hospital death rates between races was unanticipated. In 1980 the overall rates of inpa-

tient death for whites and African Americans were identical; during the study period the rates for the two races diverged. These data are consistent with data showing that African Americans tend to receive more intensive and expensive care in the final year of life.[9] These racial differences in place of death may be the result of differences in preferences, with African Americans more likely to choose life-prolonging procedures and less likely to complete advance care directives and agree to do-not-resuscitate (DNR) orders.[10] However, evidence that terminally ill African Americans receive less care from family members and friends could mean that more African Americans depend on hospitals for final care, regardless of their preferences.[11] It is also possible that African Americans and other population groups lack access to hospice and home care services in their region. Whether racial differences in care come from different preferences or disparities in access, or both, is an important question for future research.

"Whether racial differences in care come from different preferences or disparities in access is an important question for future research."

■ **Cancer.** The proportion of in-hospital deaths has fallen farther for cancer than for any other disease. In 1998 only about a third of cancer patients died as inpatients, nearly reversing the ratio of two decades earlier. Furthermore, unlike for other major causes of death, the shift for cancer has been toward residences far more than toward nursing homes. These data, combined with data indicating that as much as a half of cancer patients receive hospice care at the end of life, suggest acceptance of hospice-directed home deaths among oncologists and oncology patients.[12] It appears that special attention to end-of-life care for cancer has altered practices dramatically. One possible explanation for the relatively slight changes for other causes of death, such as COPD and heart failure, is that that impending death from these causes is less predictable. Without very reliable prognoses, it may be more difficult to stop life-prolonging treatment in favor of hospice and other kinds of out-of-hospital palliative care.

■ **Medicare spending.** Care for the dying is very expensive, consuming roughly 27 percent of Medicare's spending for the 5 percent of beneficiaries who die. It has long been hoped that such costs could be reduced by decreasing in-hospital deaths, but the percentage of Medicare spending during the last year of life has not fluctuated much even as in-hospital death has declined a great deal.[13] The steady decline in hospital deaths, combined with data suggesting no significant change in the payments for individuals who die or the proportion of Medicare expenditures going to them, indicates that reducing out-of-hospital death does not save money at the end of life.[14]

■ **Study limitations.** The principal limitation of this study is that the outcome measure is site of death, which is not necessarily the same place that the patient spent most of his or her last months. Death as a hospital inpatient does not rule out the possibility that a patient was dying at home up until the final twenty-four hours before death.

The other important limitation is in the quality of the data from the 1980s. Thirty-one percent of deceased people in that decade lived in counties where in-hospital death could not be distinguished from other facility-based death. Hence, unlike the complete 1990s data, the trend data for the 1980s are only an estimate. However, since the 1980s trend data describe just under 70 percent of U.S. deaths during the 1980s, they are probably an adequate representation.

The authors thank Kenneth Kochanek of the U.S. Centers for Disease Control and Prevention and two anonymous reviewers.

NOTES

1. See, for example, M.J. Field and C.K. Cassel, eds., *Approaching Death: Improving Care at the End of Life* (Washington: National Academies Press, 1997).
2. T. Hoyer, "A History of the Medicare Hospice Benefit,"in *A Good Dying: Shaping Health Care for the Last Months of Life*, ed. J.K. Harrold and J. Lynn (New York: Haworth Press, Inc.; 1998); and N.L. Cantor, *Legal Frontiers of Death and Dying* (Bloomington: Indiana University Press, 1987).
3. The three cases were *Cruzan by Cruzan v. Director,* Vol. 497 U.S. 261 (1990); *Washington v. Glucksberg,* Vol. 521 U.S. 702 (1997); and *Vacco, Attorney General of New York v. Quill*, Vol. 117 S.Ct. 2293 (1997). See also P. Murphy et al., "Description of the SUPPORT Intervention: Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatments," *Journal of the American Geriatrics Society* 48, no. 5 Supp. (2000): S154–S161; M.E. Baker, "Economic, Political, and Ethnic Influences on End-of-Life Decision-Making: A Decade in Review," *Journal of Health and Social Policy* 14, no. 3 (2002): 27–39; and the Education on Palliative and End-of-Life Care home page, www.epec.net (14 November 2003).
4. M.A. Sager et al., "Changes in the Location of Death after Passage of Medicare's Prospective Payment System: A National Study," *New England Journal of Medicine* 320, no. 7 (1989): 433–439; N.K. Sekhri, "Managed Care: The U.S. Experience," *Bulletin of the World Health Organization* 78, no. 6 (2000): 830–844; G.F. Anderson and P.S. Hussey, "Population Aging: A Comparison among Industrialized countries," *Health Affairs* 19, no. 3 (2000): 191–203; S.Q. Duffy and D.E. Farley, "Patterns of Decline among Inpatient Procedures," *Public Health Reports* 110, no. 6 (1995): 674–681; and K. Liu et al., *Medicare's Post-Acute Care Benefit: Background, Trends, and Issues to Be Faced*, January 1999, aspe.hhs.gov/daltcp/reports/mpacb.htm (14 November 2003).
5. Last Acts, *Means to a Better End: A Report on Dying in America Today,* November 2002, www.lastacts.org/files/misc/meansfull.pdf (14 November 2003); R.S. Pritchard 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," *Journal of the American Geriatrics Society* 46, no. 10 (1998): 1242–1250; Gallup Poll, "Knowledge and Attitudes Related to Hospice Care" (Princeton, N.J.: Gallup Organization, 1996); J. Townsend et al., "Terminal Cancer Care and Patients' Preference for Place of Death: A Prospective Study," *British Medical Journal* 301, no. 6749 (1990): 415–417; and National Hospice Organization, "Gallup Poll Reveals First American Attitudes about Terminal Illness and Hospice Care," Press Release (Arlington, Va.: National Hospice Organization, 29 July 1992).
6. Hoyer, "A History of the Medicare Hospice Benefit."
7. U.S. Department of Agriculture, Economic Research Service, "Measuring Rurality: Rural-Urban Continuity Codes," 6 August 2003, www.ers.usda.gov/briefing/rurality/RuralUrbCon (9 February 2004).
8. W.R. Levesque, "Lawsuit Alleges That Hospice Released Private Information," *St. Petersburg Times,* 2 May 2002; and R. Davis, "The Painful Truth of Hospice Care," *USA Today,* 20 August 2001.
9. E.J. Emanuel et al., "Managed Care, Hospice Use, Site of Death, and Medical Expenditures in the Last Year of Life," *Archives of Internal Medicine* 162, no. 15 (2002): 1722–1728; and C. Hogan et al., "Medicare Beneficiaries' Costs of Care in the Last Year of Life," *Health Affairs* 20, no. 4 (2001): 188–195.
10. L.A. O'Brien et al., "Nursing Home Residents' Preferences for Life-Sustaining Treatments," *Journal of the American Medical Association* 274, no. 22 (1995): 1775–1779; and L.B. Shepardson et al., "Racial Variation in the Use of Do-Not-Resuscitate Orders," *Journal of General Internal Medicine* 14, no. 1 (1999): 15–20.
11. E.J. Emanuel et al., "Assistance from Family Members, Friends, Paid Care Givers, and Volunteers in the Care of Terminally Ill Patients," *New England Journal of Medicine* 341, no. 13 (1999): 956–963.
12. M. Agrawal and E.J. Emanuel, "Attending to Psychologic Symptoms and Palliative Care," *Journal of Clinical Oncology* 20, no. 3 (2002): 624–626.
13. J.D. Lubitz and G.F. Riley, "Trends in Medicare Payments in the Last Year of Life," *New England Journal of Medicine* 328, no. 15 (1993): 1092–1096; and P.A. Singer and F.H. Lowy, "Rationing, Patient Preferences, and Cost of Care at the End of Life," *Archives of Internal Medicine* 152, no. 3 (1992): 478–480.
14. E.J. Emanuel and L.L. Emanuel, "The Economics of Dying: The Illusion of Cost Savings at the End of Life," *New England Journal of Medicine* 330, no. 8 (1994): 540–544.

1062-8606/04/28–36$03.00/0
AMERICAN JOURNAL OF MEDICAL QUALITY

Vol 19, No 1
Jan/Feb 2004

Depression Management in Medical Clinics: Does Healthcare Sector Make a Difference?

Susan M. Frayne, MD, MPH, Karen M. Freund, MD, MPH, Katherine M. Skinner, PhD, Arlene S. Ash, PhD, and Mark A. Moskowitz, MD

Medical providers often fail to treat depression. We examined whether treatment is more aggressive in a setting with accessible mental health resources, the Veterans Health Administration (VA). VA and non-VA primary care physicians and medical specialists viewed a videotape vignette portraying a patient meeting criteria for major depression and then answered interviewer-administered questions about management. We found that 24% of VA versus 15% of non-VA physicians would initiate guideline-recommended treatment (antidepressants or mental health referral, or both) ($P = .09$). Among those who identified depression as likely, 42% of VA versus 19% of non-VA physicians would treat ($P = .002$): 23% versus 3% recommended mental health referral ($P < .001$) and 21% versus 17% an antidepressant ($P = .67$). Although many patients with major depression may not receive guideline-recommended management, VA physicians do initiate mental health referral more often than do non-VA physicians. Access to mental health services may prove valuable in the campaign to increase physician adherence to depression clinical guidelines.

Key words: Depression; models, organizational; physician's practice patterns; primary health care; quality assurance, health care; veterans.

Dr Frayne was Staff Physician, VA Boston Healthcare System, Associate Professor of Medicine and Public Health, Boston University, Boston, Mass, and Research Associate, Center for Health Quality, Outcomes, and Economic Research, VA Medical Center, Bedford, Mass; Dr Freund is Director of Women's Health Unit, Evans Department of Medicine, Boston Medical Center and Professor of Medicine, Boston University, Boston, Mass; Dr Skinner is Associate Professor of Public Health, Boston University, Boston, Mass, and Research Associate, Center for Health Quality, Outcomes and Economic Research, VA Medical Center, Bedford, Mass; Dr Ash is Research Professor of Medicine, Boston University, Boston, Mass.

Dr Moskowitz was Chief of General Internal Medicine and Professor of Medicine, Boston University, Boston, Mass (Deceased).

Portions of this work were presented at the VA Health Services Research & Development National Meeting, Washington, DC, February 2001, and at the Society of General Internal Medicine National Meeting, San Diego, Calif, May 2001.

The views expressed herein are those of the authors and do not necessarily reflect the views of the Department of Veterans Affairs.

Corresponding author: Susan M. Frayne, MD, MPH, Center for Health Care Evaluation, VA Palo Alto, 795 Willow Road (152-MPD), Menlo Park, CA 94025 (e-mail: sfrayne@stanford.edu).

Clinicians, health services researchers, and policymakers alike have been struck by mounting evidence of institution-level variability in the quality of care provided for many health conditions. Organizational characteristics such as funding mechanisms and mission appear to influence clinical management (1–4). One major organizational model is the Veterans Health Administration (VA), the largest integrated health care system in the United States, serving 3.7 million patients with a congressionally appropriated budget of almost $17 billion in 2000 (5). Recent studies have examined differences in VA versus private sector care (6–8), but institutional variability in care for major depression has received less attention. Documenting differences in care delivery is the first step toward achieving convergence to a higher quality standard.

Major depression is a promising index condition for comparing quality of ambulatory care provided in VA versus the private sector. Depression is common, affecting 5–10% of patients in medical ambulatory care settings (9), where it typically presents (10). Rates are even higher in a geriatric population (11, 12). Depression is highly morbid (13–15). Treatment improves outcomes (16) and can be provided at relatively modest cost (17). However, medical providers in the private sector markedly undertreat depression (18, 19), despite the availability of Agency for Healthcare Research and Quality (AHRQ) clinical guidelines encouraging clinicians to respond aggressively to depression with antidepressant medications or mental health referral (or both) (rather than waiting for alternative diagnoses to be excluded) and to see the patient in follow-up within 1–2 weeks after identifying depression (20). We hypothesized that the greater availability of mental

health services in VA compared with the private sector (21–25) should result in better adherence to depression management guidelines in a VA setting.

This hypothesis is difficult to test because the case mix of patients in the 2 sectors is so different (26). We used an experimental design to overcome this difficulty. We compared the management recommendations of VA versus non-VA physicians who viewed videotapes portraying a standardized case of major depression in an elderly patient.

METHODS

Experimental Design Overview

Physicians practicing in an outpatient VA or non-VA medical setting in the northeastern United States were randomly selected to view a professionally acted videotape vignette portraying an elderly patient who meets criteria for a diagnosis of major depression. After viewing the videotape, each participant answered interviewer-administered, open-ended questions about his or her differential diagnosis, treatment recommendations, and follow-up recommendations. We then compared the management recommendations of VA versus non-VA physicians. Of note, AHRQ depression management guidelines had been disseminated nationally at least 3 years before the first interview.

Physician Subjects

Institutional Review Board approval was obtained from New England Research Institute, from Boston University, and from each of the 10 participating VA facilities; subjects provided written informed consent. We randomly selected New England internists and family practitioners who provide care in the ambulatory setting; recruitment was extended to New York to increase the pool of eligible physician subjects with characteristics of interest. Our objective was to identify a representative cohort of physicians providing ambulatory primary care services from each setting (VA and non-VA), stratified on physician characteristics of interest.

For the non-VA cohort, randomly selected physicians identified from Board of Registration in Medicine listings were recruited from 1996 to 1997. Physicians were eligible for inclusion if they reported involvement in patient care for at least 8 hours per week with at least 60% of their practice being primary care. Female and family practice physicians were oversampled, as were African-American physicians (using a snowball recruitment technique [27]) to meet additional study objectives not pursued here.

With recruitment of non-VA subjects ongoing, we initiated a second study phase to replicate the non-VA design in a VA setting, thereby allowing for direct VA–non-VA comparisons. A random sample of VA staff physicians who provide primary care or medical subspecialty care (or both) was identified from lists provided by the Department of Medicine/Ambulatory Care; VA subjects were recruited from 1999 to 2001. Medical subspecialists were specifically included in the VA study because a substantial proportion of VA primary care is delivered in specialty clinics. A physician was defined to be a specialist if he or she reported providing any specialty care. Female and medical subspecialist physicians were oversampled; because of the smaller sampling frame, it was not feasible to oversample African-American physicians.

After sending an introductory letter describing a study about "physicians' clinical decision-making," study staff recruited subjects by telephone. Participating subjects received $100 or its equivalent in educational materials.

Interview Procedures

The standardized interview protocol was revised after field testing to optimize its validity. After showing the physician-subject the depression vignette, a trained interviewer asked physicians to assign a probability to each condition listed in their differential diagnosis, then indicate what tests, referrals, or medications (or all) they would order, and how soon they would see the patient again. At the end of the interview, subjects provided personal demographic information and characterized their clinical practice. The interview took approximately 1 hour and was audiotaped. After primary coding by a research assistant, a 20% random sample was rereviewed by the investigators for coding quality control.

The interview was conducted at the physician's practice site during regular office hours. Physicians were not told that the patient and the physician in the videotape were actors. They were instructed to use their own practice as the frame of reference for their management recommendations (VA physicians with private practices were instructed to use their VA practice).

Instruments: Videotaped Vignettes

A 5-minute, videotaped clinical vignette portraying a patient meeting full diagnostic criteria for late life,

major depression was professionally produced. In the vignette, the patient presents to a physician for a hypertension follow-up visit, complaining of constipation. During the course of the interview, the patient describes specific symptoms meeting criteria (28) for depression: anhedonia, anorexia, weight loss, insomnia, fatigue, psychomotor retardation, and thoughts of death. The script did not include specific mention of depressed mood or suicidal ideation for 2 reasons. First, late-life depression often presents with dominant somatic symptoms (11). Second, we wished to evaluate the variability in subjects' responses ie, *differences* in VA versus non-VA care; therefore, we developed a case where a broader differential diagnosis was possible.

Substantial efforts were made to enhance the realism of the vignette. The script was derived from actual primary care cases; transcripts from an audiotaped session in which clinicians role-played these cases served as a basis for script development. An effort was made to use common lay terms (eg, "I just don't seem to have the pep I used to"). Three panels of practicing physicians confirmed the script's authenticity. During filming, an internist with experience in managing depression assured that nonverbal cues were invariant across all versions of the videotape and were consistent with the clinical presentation of major depression.

To meet study objectives not discussed here (29, 30), patient characteristics (including age, which was either 67 or 79 years) were systematically varied in different versions of the video. However, the clinical content and nonverbal communication in the single depression vignette viewed by each physician were held constant.

Outcome Measures and Covariates

The dependent variables were drawn from physicians' responses to open-ended questions about treatment recommendations and monitoring. A physician was considered to have recommended guideline-recommended (20, 31) treatment if he or she recommended starting an antidepressant or referral to a mental health practitioner (or both) at the initial visit. Monitoring occurred if the provider recommended follow-up within the AHRQ-recommended interval of 2 weeks. Physicians were considered to have identified depression if, in response to the open-ended question about differential diagnosis, they said the patient portrayed in the video more likely than not (probability $\geq$50%) had depression.

Independent measures included characteristics of the physician-subjects: sex, age, years since medical school graduation, whether he or she provided any medical subspecialty care, whether he or she provided any primary care, and whether he or she was board certified. Practice characteristics examined were self-reported number of outpatients seen per week and time allotted to a follow-up patient. Caseload characteristics were the physicians' estimates of what percent of their caseloads were female, Caucasian, and aged 65 years or older. Patient characteristics systematically varied across videotape versions were sex, race, and age.

Analysis

In bivariate analyses, we compared VA versus non-VA physicians on demographic and practice characteristics. We next compared the 2 groups of physicians on the dichotomous outcomes of interest (treatment and monitoring). For descriptive purposes, we also characterized their specific treatment recommendations (antidepressant use, mental health referral). We then conducted logistic regression analyses to assess the effect of institution (VA versus non-VA) on our 2 main outcomes (treatment and monitoring), controlling for characteristics of the "patient" portrayed in the videotape (sex, race, and age) and characteristics of the physician-subject (sex and specialty status).

We next repeated the same analyses on the subset who identified depression in the vignette to see whether those physicians who recognized depression took action to address it. In bivariate analyses, we then examined treatment, monitoring, and specific treatment recommendations in this subset. We also conducted logistic regression analyses on this subset to assess the effect of institution (VA versus non-VA) on our 2 main outcomes (treatment and monitoring), controlling for patient and physician characteristics.

After completing these main analyses with the full cohort and with the subset who identified depression, we examined other specific subsets of our full cohort, again comparing VA versus non-VA physicians. Because AHRQ recommends that patients started on an antidepressant should be seen within 2 weeks to assess the effect of therapy, we examined monitoring recommendations in the subset who recommended an antidepressant. Because providers may have recommended treatment at a follow-up visit, we also examined recommended patient monitoring in the subset of subjects who recommended neither an antidepressant nor a referral.

We conducted various sensitivity analyses. First, we repeated all analyses using an expanded interpretation of the AHRQ standard, where the definition of

Table 1

Characteristics of VA and Non-VA Study Physicians and Their Practices[a]

	VA (N = 115)	Non-VA (N = 128)	P Value[b]
Physician characteristics			
Male (%)	53.0	62.5	.14
Age, mean (SD) (y)	47.1 (9.1)	45.2 (9.7)	.13
Years since medical school graduation, mean (SD)	20.1 (9.6)	18.0 (9.9)	.11
Does any specialty work (%)	49.6	25.8	<.001
Does any primary care (%)	77.4	100	<.001
Board certified (%)	89.6	100	<.001
Practice characteristics			
Number of outpatients seen per week, mean (SD)	38.2 (26.6)	75.0 (53.3)	<.001
Minutes per follow-up patient, mean (SD)	23.7 (5.8)	16.9 (5.0)	<.001
Percentage of patients female, mean (SD)	8.9 (13.8)	62.3 (15.0)	<.001
Percentage of patients Caucasian, mean (SD)	79.5 (16.6)	69.0 (32.4)	.002
Percentage of patients ≤65 y, mean (SD)	30.8 (16.7)	74.0 (19.1)	<.001

[a] VA indicates Veterans Health Administration; SD, standard deviation.
[b] Statistical significance for comparison of VA versus non-VA physicians.

treatment included recommending follow-up within 2 weeks (if depression was identified). The rationale for this is that some physicians who identify depression as a possibility will plan to see the patient back within 2 weeks and begin depression treatment at that follow-up visit. We also expanded the definition of adequate monitoring from 2 weeks to 1 month, longer than the AHRQ-recommended interval (20). Second, to ensure that differences in the distribution of specialists versus generalists in the 2 samples were not driving our findings, we repeated all analyses, limiting the VA sample to only those who perform at least some primary care work.

Using chi-square tests for discrete variables and t tests for continuous variables, we calculated P values for univariate comparisons, with $P < .05$ considered statistically significant. We calculated odds ratios and 95% confidence intervals for multivariable outcomes. The SAS statistical computing package was used for all analyses. A distinct advantage of the study's fractional factorial design (described previously [29]) was that it made it possible to use a relatively small sample to efficiently examine the independent effects of predetermined patient-level, provider-level, and institution-level factors.

RESULTS

In the VA sample, 78% of contacted physicians participated (N = 115). In the non-VA sample, the response rate before the snowball recruitment phase was 85% (N = 128). Among VA physicians, 27% also had private practices. Among non-VA physicians, none had VA practices, and 98% cared for at least some patients under managed care contracts.

Demographic and Practice Characteristics of Subjects

As Table 1 shows, VA and non-VA study physicians were similar in sex, age, and years since medical school graduation. Differences in the proportions providing specialty and primary care reflect differences in the sampling strategy rather than differences in the distribution of generalists and specialists in these 2 institutional settings.

Practice characteristics differed substantially between the VA and non-VA groups (Table 1). In contrast to the non-VA physicians, VA physicians described their practices as predominantly composed of older, male patients.

Treatment and Monitoring of Depression

Table 2 compares VA and non-VA physicians' treatment and monitoring recommendations for the depression vignette. We found that 24% of VA versus 15% of non-VA physicians recommended treatment (P = .09) and 32% versus 53% recommended adequate monitoring ($P < .001$).

Fifty-four percent of VA versus 73% of non-VA physicians (P = .002) considered the likelihood of depression in the vignette patient to be ≥50%. Thirty-four percent of VA and 20% of non-VA physicians listed 5 or more nondepression diagnoses on their differential

Table 2

Treatment and Monitoring of Depression by VA[a] and Non-VA Study Physicians Among the Full Cohort and Among the Subset Who Identified Depression

	Full Cohort			Subset Who Identified Depression		
	VA (N = 115)	Non-VA (N = 128)	P Value[b]	VA (N = 62)	Non-VA (N = 93)	P Value[b]
Recommend treatment (%)[c]	23.5	14.8	.09	41.9	19.4	.002
Start antidepressant (%)	11.3	13.3	.70	21.0	17.2	.67
Mental health referral (%)	13.0	2.3	.002	22.6	3.2	<.001
See within 2 w (%)[d]	32.2	53.1	<.001	37.1	55.9	.02

[a] VA indicates Veterans Health Administration.
[b] Statistical significance: comparison of VA versus non-VA physicians.
[c] Recommended starting antidepressant or referral to mental health (or both).
[d] Recommended follow-up within 2 weeks.

diagnosis (P = .01). Among the subset who identified depression, 42% of VA and 19% of non-VA physicians recommended treatment (P = .002). The difference between VA and non-VA physicians primarily reflects higher rates of recommending a mental health referral rather than major differences in antidepressant recommendation: 23% versus 3% of the subset who identified depression recommended mental health referral (P < .001).

The next analyses focused on other subsets of the full study group. Among those who recommended starting antidepressant therapy, 46% of VA versus 41% of non-VA doctors said they would see the patient back within 2 weeks (P = .79). Among those who did not recommend any type of treatment (antidepressants or referral), 30% of VA and 56% of non-VA physicians said they would see the patient back within 2 weeks (P < .001).

The results of our multivariable analyses are presented in Table 3. After adjusting for patient characteristics (sex, race, and age) and physician characteristics (sex and specialty status), VA physicians were twice as likely to recommend treatment and half as likely to recommend monitoring within 2 weeks, compared with non-VA physicians. The positive effect of VA status on treatment was even more pronounced in the subset who identified depression.

Table 3

Differences in Depression Guideline Adherence by Health Care Sector

	Full Cohort[a]		Subset Who Identified Depression	
	OR[b]	95% CI	OR[b]	95% CI
Recommend Treatment[c]	2.0	1.0–4.0	3.8	1.7–8.5
See within 2 w[d]	0.4	0.3–0.7	0.4	0.2–0.9

[a] OR indicates odds ratio; CI, confidence interval; VA, Veterans Health Administration.
[b] Odds Ratio for comparison of VA versus non-VA physicians, adjusted for patient characteristics (sex, race, and age) and physician characteristics (sex and specialty status).
[c] Recommended starting antidepressant or referral to mental health (or both).
[d] Recommended follow-up within 2 weeks.

Sensitivity Analyses

The first set of sensitivity analyses used the expanded definitions of guideline adherence. When the definition of "treatment" is expanded to include follow-up within 2 weeks, VA physicians were less likely than non-VA physicians to recommend treatment (33% versus 49%, P = .01). The direction of the association between institution and treatment reverses in this sensitivity analysis, reflecting the fact that VA physicians were substantially less likely to recommend follow-up within 2 weeks. VA physicians were also less likely than non-VA physicians to recommend follow-up within the longer interval of 1 month (78% versus 91%, P = .006).

In the second set of sensitivity analyses, we limited the VA sample to only those physicians who provide primary care to at least some of their patients (N = 89). In this subset, 35% of VA versus 26% of non-VA physicians provided any specialty care. Our findings were unchanged when we repeated the VA–non-VA comparisons using this subset. For example, comparing VA versus non-VA physicians, 25% versus 15% recommended treatment (P = .07), 13% versus 13% recommended an antidepressant (P = .94), 13% versus 2% recommended mental health referral (P = .002), and

30% versus 53% requested follow-up within 2 weeks ($P < .001$).

CONCLUSIONS

When presented with a videotaped patient scenario meeting criteria for late-life major depression, VA physicians providing ambulatory medical care—like the non-VA physicians with whom they were compared—often did not recommend responding aggressively with early treatment interventions and close follow-up. We did, however, identify specific institution-level differences: VA physicians were substantially more likely than non-VA physicians to recommend referral to mental health and less likely to recommend follow-up within the AHRQ-recommended interval of 2 weeks. Our findings point to possible systems-level explanations for physicians' nonintervention.

Less than 1 in 4 physicians in either setting indicated that they would have taken any steps to treat the patient in the video, who met criteria for major depressive disorder. If these responses reflect actual practice, more than 3 of 4 depressed older patients presenting to a study physician would not receive early intervention. On the basis of earlier work (32), even those treated with antidepressants might have received subtherapeutic doses. Although studies examining physicians' self-reported practice style have suggested higher levels of adherence to depression treatment guidelines than we saw (3, 33), most (34–36), though not all (37), studies that directly observe actual practice behavior have had findings more in line with ours.

The follow-up management of over half of the physician-subjects likewise was not consistent with AHRQ guidelines: only 32% of VA and 53% of non-VA physicians indicated that they would see the patient depicted for a 2-week follow-up visit. Even this is an optimistic estimate of physician monitoring because physicians who did ask for a 2-week follow-up visit were not necessarily planning to pursue depression; some intended to explore a competing consideration in the differential diagnosis. Although a parallel process of medical work-up and depression treatment may be indicated, such delay could have serious consequences, especially for a patient like the one depicted in the videotape who had already suffered depression's adverse effects on quality of life (13) for 2 months before presenting for care.

The literature on quality of depression care has emphasized identification and treatment of depression; monitoring has received less attention. A particular advantage of our study design is that we were able not only to characterize monitoring decisions but also to link treatment recommendations to monitoring plans. We found that the subgroup who did recommend treatment often did not recommend monitoring of the therapy closely. Among those who recommended starting an antidepressant, over half did not ask to see the patient again within the AHRQ-recommended 2-week interval. If antidepressant therapy is initiated without monitoring, adverse effects of therapy—including discontinuation of therapy, side effects, and suicidality (25, 38)—may go unrecognized.

Although our study was not designed to determine the causes of nonintervention, it provides clues about some of the competing explanations. First, our findings argue against the hypothesis that failure to treat at an initial visit typically reflects a plan for watchful waiting. Among those physicians who did not recommend treatment, only 30% of VA and 56% of non-VA physicians recommended follow-up within the AHRQ-recommended interval of 2 weeks. We did not examine what action physicians would recommend at the follow-up visit.

Second, our findings refute the hypothesis that failure to treat is solely a result of failure to recognize depression. Even among those who thought the patient depicted probably had depression, a minority of physicians recommended treatment with antidepressants or mental health referral (or both). However, the fact that VA providers were substantially more likely to recommend treatment if they recognized depression suggests underrecognition of depression as 1 factor explaining undertreatment in VA. Although it is surprising that we saw lower rates of recognition of depression in VA than in the private sector, this may reflect physicians' Bayesian thinking in response to case-mix differences between the 2 settings. Despite its high prevalence in VA (9, 39–41), depression in VA frequently presents with other psychiatric comorbidities (42, 43); therefore, VA providers may be less accustomed to recognizing isolated depression, as depicted in the vignette. Alternatively, the fact that VA providers care for a medically sicker patient population (6, 44) may have led them to favor competing explanations for the patient's somatic symptoms (eg, cancer) (45, 46). VA doctors were more likely than non-VA doctors to list a large number of medical conditions as potential explanations for the patient's symptoms, supporting this possibility.

Third, the VA–non-VA differences we observed are consistent with our hypothesis that systems factors contribute to the observed failure to intervene. On the one hand, follow-up was less aggressive in VA than in

the private sector. This might reflect disincentives inherent in a capitated system or an altered perception of what constitutes a "usual" follow-up interval for an isolated condition (where a typical patient has multiple conditions).

On the other hand, we found that VA physicians were more likely than non-VA practitioners to recommend referral of a patient with depression to a mental health provider. This may reflect access to mental health services. In VA, primary care providers can refer directly to mental health clinics (usually on-site). Many VA facilities have a mental health triage department, providing a central access point. In contrast, in the private sector, under coverage, copays or caps to the number of mental health visits allowed by the payer may limit access (22, 24, 25). Although our study was not designed to determine whether mental health referral leads to better patient outcomes, mental health referral does qualify as guideline-concordant care (20).

There are other possible explanations for the VA–non-VA differences in mental health referral. There may be less stigma associated with a depression diagnosis in VA, where mental illness is more prevalent (39). Pressured by the competing demands of multiple medical and social problems (6, 44, 47), and largely free of the disincentives to referral seen in managed care (1, 48, 49), VA physicians may refer to limit the scope of their responsibility. However, the possibility that our findings about observed decision-making behavior do indeed reflect systems issues is supported by earlier work in which self-reported behaviors were examined across systems of care. Meredith and colleagues found that physicians practicing in staff/group model managed care organizations (including VA) were more likely than physicians practicing in network model managed care organizations to report referring patients with depression to a mental health provider and less likely to report starting an antidepressant themselves (3).

This study has several limitations. First is the issue of external validity, ie, the ability of a vignette to elicit responses reflecting actual practice (50). However, a major study found that written vignettes provide valid estimates of process of care (50); given their greater realism, video vignettes like ours would be expected to perform even better. Indeed, our study design (which included developing a realistic vignette, interviewing the physician in his or her practice context, and asking the physician to use his or her own practice as the frame of reference for responses) appears to have been successful: participants made spontaneous comments like, "I saw that patient this morning," suggesting that they saw the vignette as a real clinical encounter. Our earlier work with this videotape methodology indicated that 88% of physicians reported that the case depicted in the videotape was typical of their own practice (29).

A second limitation is that we examined management decisions for a single visit rather than longitudinally over an episode of care (51). Although this gives a helpful picture of the initial response, it is possible that many physicians reassess patients such as this within a short interval, initiating treatment at the second visit. However, the fact that 70% of VA and 44% of non-VA physicians who did not recommend treatment for depression also did not plan to see the patient back within a 2-week interval argues against this.

Third, the VA sample was enriched with medical subspecialists, by design, to more accurately reflect the physician population providing primary care to depressed patients in VA. This raises the possibility that the observed VA–non-VA differences actually represent generalist-specialist differences, which are known to influence management of several conditions (33, 52, 53). However, when we repeated analyses limiting the VA sample to only those physicians who spend at least part of their time in primary care, our findings were unchanged, suggesting that the observed VA–non-VA differences are not an artifact of our selection criteria. Another potential selection issue is that VA interviews started 2 years after the non-VA interviews ended. Although AHRQ guidelines had been nationally disseminated 3 years before the first non-VA interview, it is unknown whether any secular trends could have influenced our findings. However, by the time our study started, about 60% of physicians responding to 1 survey reported familiarity with the guidelines (54).

Fourth, our findings might not generalize to physicians practicing outside the northeastern United States. Regional practice variation has been documented previously (55, 56).

Our study design also had distinct advantages. First, we used an experimental methodology in which all physicians viewed an identical clinical scenario. This allowed for direct comparisons between physicians practicing in very different clinical settings, overcoming the problem of marked variability in patient case mix across systems of care. Second, the open-ended format of our questions allowed physicians to report the diagnoses and management approaches consistent with their own practice styles (50, 57). Third, we were able to link treatment recommendations with follow-up plans, to characterize management as a whole. Our study augments what is currently a very small body of literature about the degree to which financing mech-

anisms and organization of health care may influence delivery of mental health services.

In conclusion, on the basis of physician responses to a 5-minute videotaped vignette depicting an elderly patient with major depression, in both VA and the private sector, the degree to which physicians' recommended interventions correspond with standard treatment and monitoring guidelines is low. However, in VA, which emphasizes mental health services, physicians are substantially more likely to recommend referral to mental health than are their private-sector counterparts. Our findings have 3 major implications. First, many elderly VA and private sector patients with major depressive disorder may not be receiving treatment at initial presentation. This has a huge human cost: untreated depression leads to substantial decrements in quality of life and functional status (13–15). This also has an important economic cost to the health care system and to society at large (58). Second, interventions designed to decrease the undertreatment of depression in primary care need to be tailored to institutional setting. Underrecognition of major depression had a greater impact on treatment decisions in VA, so interventions aimed at increasing recognition of depression may still have value in VA. In contrast, nontreatment after successful recognition of depression was the dominant scenario in the non-VA setting; alternate interventions would be needed there. Third, models of health care delivery in which mental health care services are readily accessible to patients (as in VA) deserve more attention. Increased access to mental health services can be achieved at relatively modest cost (17) and can prevent substantial psychiatric morbidity (24, 59). Assuring parity in access to mental health versus other routine health care services in the private sector and avoiding erosion of access to mental health services in the VA setting may prove important if we are to see continued improvement in physician adherence to clinical guidelines for depression, the second leading cause of disability worldwide (13).

ACKNOWLEDGMENTS

This study was supported by Department of Veterans Affairs Health Services R&D grants GEN 970231-1 and RCD 98-312, and National Institute on Aging grant AG12437-03. John B. McKinlay, PhD, of New England Research Institute initially developed the research methodology used in this study; he along with Julie T. Irish, PhD, and Ting H. Lin, PhD, were central to the design and execution of the first phase of this study. B. Graeme Fincke, MD, and Dan Berlowitz, MD, MPH, provided extremely valuable editorial suggestions on earlier versions of this manuscript. The authors are grateful for the extensive technical assistance provided by Stephanie J. Smith, Laurie Craigen, Jeff Murphy, Patricia Loya, Danielle Pierce, and Hai Lin, PhD.

References

1. Pantilat SZ, Chesney M, Lo B. Effect of incentives on the use of indicated services in managed care. *West J Med.* 1999;170:137–142.
2. Greenfield S, Nelson EC, Zubkoff M, et al. Variations in resource utilization among medical specialties and systems of care. Results from the medical outcomes study. *JAMA.* 1992;267:1624–1630.
3. Meredith LS, Rubenstein LV, Rost K, et al. Treating depression in staff-model versus network-model managed care organizations. *J Gen Intern Med.* 1999;14:39–48.
4. Allison JJ, Kiefe CI, Weissman NW, et al. Relationship of hospital teaching status with quality of care and mortality for Medicare patients with acute MI. *JAMA.* 2000;284:1256–1262.
5. Allocation Resource Center. Fiscal Year 2000 Funding Allocations—Final Data Release—December 20, 1999; VERA 2000 Information and Reports; Veterans Health Administration. Available at: vaww.arc.med.va.gov.
6. Petersen LA, Normand SL, Daley J, et al. Outcome of myocardial infarction in Veterans Health Administration patients as compared with medicare patients. *N Engl J Med.* 2000;343:1934–1941.
7. Jha AK, Shlipak MG, Hosmer W, et al. Racial differences in mortality among men hospitalized in the Veterans Affairs health care system. *JAMA.* 2001;285:297–303.
8. Peabody JW, Luck J. How far down the managed care road? A comparison of primary care outpatient services in a Veterans Affairs medical center and a capitated multispecialty group practice. *Arch Intern Med.* 1998;158:2291–2299.
9. Katon W, Schulberg H. Epidemiology of depression in primary care. *Gen Hosp Psychiatry.* 1992;14:237–247.
10. Regier DA, Goldberg ID, Taube CA. The de facto US mental health services system: a public health perspective. *Arch Gen Psychiatry.* 1978;35:685–693.
11. Stewart RB, Blashfield R, Hale WE, et al. Correlates of Beck Depression Inventory scores in an ambulatory elderly population: symptoms, diseases, laboratory values, and medications. *J Fam Pract.* 1991;32:497–502.
12. 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.
13. Murray CJ, Lopez AD. Alternative projections of mortality and disability by cause 1990–2020: Global Burden of Disease Study. *Lancet.* 1997;349:1498–1504.
14. Callahan CM, Wolinsky FD, Stump TE, et al. Mortality, symptoms, and functional impairment in late-life depression. *J Gen Intern Med.* 1998;13:746–752.
15. Covinsky KE, Fortinsky RH, Palmer RM, et al. Relation between symptoms of depression and health status outcomes in acutely ill hospitalized older persons. *Ann Intern Med.* 1997;126:417–425.
16. Coulehan JL, Schulberg HC, Block MR, et al. Treating depressed primary care patients improves their physical, mental, and social functioning. *Arch Intern Med.* 1997;157:1113–1120.
17. Sturm R. How expensive is unlimited mental health care coverage under managed care? *JAMA.* 1997;278:1533–1537.
18. Wells KB. Depression in general medical settings. Implications of three health policy studies for consultation-liaison psychiatry. *Psychosomatics.* 1994;35:279–296.

19. Freund KM, Moskowitz MA, Lin TH, et al. Early antidepressant therapy for elderly patients. *Am J Med.* 2003;114:15–19.
20. Depression Guideline Panel. *Depression in Primary Care. Vol 2. Treatment of Major Depression. Clinical Practice Guideline, Number 5.* Rockville Md: US Department of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research; 1993. AHCPR Publication 93-0551.
21. Wells KB, Sherbourne C, Schoenbaum M, et al. Impact of disseminating quality improvement programs for depression in managed primary care: a randomized controlled trial. *JAMA.* 2000;283:212–220.
22. Simon GE, Grothaus L, Durham ML, et al. Impact of visit copayments on outpatient mental health utilization by members of a health maintenance organization. *Am J Psychiatry.* 1996; 153:331–338.
23. Jones LR, Parlour RR, Badger LW. Mental health services in primary health care: yet another era of disillusionment? *Med Care.* 1982;20:1222–1232.
24. Soumerai SB, McLaughlin TJ, Ross-Degnan D, et al. Effects of a limit on Medicaid drug-reimbursement benefits on the use of psychotropic agents and acute mental health services by patients with schizophrenia. *N Engl J Med.* 1994;331:650–655.
25. 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.
26. Iezzoni L. *Risk Adjustment for Measuring Healthcare Outcomes.* 2nd ed. Chicago: Health Administration Press; 1997.
27. Biernacki P, Waldorf D. Snowball sampling: problems and techniques of chain referral sampling. *Sociological Methods Res.* 1981;10:141–163.
28. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders.* 4th ed. Washington, DC: American Psychiatric Press; 1994.
29. Feldman HA, McKinlay JB, Potter DA, et al. Nonmedical influences on medical decision making: an experimental technique using videotapes, factorial design, and survey sampling. *Health Serv Res.* 1997;32:343–366.
30. McKinlay JB, Lin T, Freund K, et al. The unexpected influence of physician attributes on clinical decisions: results of an experiment. *J Health Soc Behav.* 2002;43:92–106.
31. Depression Guideline Panel. *Depression in Primary Care. Vol 1. Detection and Diagnosis. Clinical Practice Guideline Number 5.* Rockville Md: US Department of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research; 1993. AHCPR Publication 93-0550.
32. Rost K, Williams C, Wherry J, et al. The process and outcomes of care for major depression in rural family practice settings. *J Rural Health.* 1995;11:114–121.
33. Williams JW Jr, Rost K, Dietrich AJ, et al. Primary care physicians' approach to depressive disorders. Effects of physician specialty and practice structure. *Arch Fam Med.* 1999;8:58–67.
34. van Os TW, Ormel J, van den Brink RH, et al. Training primary care physicians improves the management of depression. *Gen Hosp Psychiatry.* 1999;21:168–176.
35. Callahan CM, Hendrie HC, Dittus RS, et al. Improving treatment of late life depression in primary care: a randomized clinical trial. *J Am Geriatr Soc.* 1994;42:839–846.
36. Callahan CM, Dittus RS, Tierney WM. Primary care physicians' medical decision making for late-life depression. *J Gen Intern Med.* 1996;11:218–225.
37. Carney PA, Dietrich AJ, Eliassen MS, et al. Recognizing and managing depression in primary care: a standardized patient study. *J Fam Pract.* 1999;48:965–972.
38. Cooper-Patrick L, Powe NR, Jenckes MW, et al. Identification of patient attitudes and preferences regarding treatment of depression. *J Gen Intern Med.* 1997;12:431–438.
39. Hankin CS, Spiro A III, Miller DR, et al. Mental disorders and mental health treatment among US Department of Veterans Affairs outpatients: the Veterans Health Study. *Am J Psychiatry.* 1999;156:1924–1930.
40. Rosenheck R, DiLella D. *Department of Veterans Affairs National Mental Health Program Performance Monitoring System: Fiscal Year 1999 Report (Draft).* West Haven: Northeast Program Evaluation Center, VA Connecticut Healthcare System; February 2, 2000.
41. 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.
42. Moos RH, Mertens JR. Patterns of diagnoses, comorbidities, and treatment in late-middle-aged and older affective disorder patients: comparison of mental health and medical sectors. *J Am Geriatr Soc.* 1996;44:682–688.
43. Friedman MJ, Schnurr PP, McDonagh-Coyle A. Post-traumatic stress disorder in the military veteran. *Psychiatr Clin North Am.* 1994;17:265–277.
44. Kazis LE, Miller DR, Clark J, et al. Health-related quality of life in patients served by the Department of Veterans Affairs: results from the Veterans Health Study. *Arch Intern Med.* 1998; 158:626–632.
45. Rost K, Nutting P, Smith J, et al. The role of competing demands in the treatment provided primary care patients with major depression. *Arch Fam Med.* 2000;9:150–154.
46. Klinkman MS. Competing demands in psychosocial care. A model for the identification and treatment of depressive disorders in primary care. *Gen Hosp Psychiatry.* 1997;19:98–111.
47. Wilson NJ, Kizer KW. The VA health care system: an unrecognized national safety net. *Health Aff (Millwood).* 1997;16:200–204.
48. Mitchell JM, Hadley J, Sulmasy DP, et al. Measuring the effects of managed care on physicians' perceptions of their personal financial incentives. *Inquiry.* 2000;37:134–145.
49. Grumbach K, Osmond D, Vranizan K, et al. Primary care physicians' experience of financial incentives in managed-care systems. *N Engl J Med.* 1998;339:1516–1521.
50. Peabody JW, Luck J, Glassman P, et al. Comparison of vignettes, standardized patients, and chart abstraction: a prospective validation study of 3 methods for measuring quality. *JAMA.* 2000;283:1715–1722.
51. Hornbrook MC, Hurtado AV, Johnson RE. Health care episodes: definition, measurement and use. *Med Care Rev.* 1985;42:163–218.
52. Ayanian JZ, Hauptman PJ, Guadagnoli E, et al. Knowledge and practices of generalist and specialist physicians regarding drug therapy for acute myocardial infarction. *N Engl J Med.* 1994; 331:1136–1142.
53. Meredith LS, Wells KB, Camp P. Clinician specialty and treatment style for depressed outpatients with and without medical comorbidities. *Arch Fam Med.* 1994;3:1065–1072.
54. Banazak DA. Late-life depression in primary care. How well are we doing? *J Gen Intern Med.* 1996;11:163–167.
55. Cowper PA, DeLong ER, Peterson ED, et al. Geographic variation in resource use for coronary artery bypass surgery. IHD Port Investigators. *Med Care.* 1997;35:320–333.
56. Ashton CM, Petersen NJ, Souchek J, et al. Geographic variations in utilization rates in Veterans Affairs hospitals and clinics. *N Engl J Med.* 1999;340:32–39.
57. Sandvik H. Criterion validity of responses to patient vignettes: an analysis based on management of female urinary incontinence. *Fam Med.* 1995;27:388–392.
58. Greenberg PE, Stiglin LE, Finkelstein SN, et al. The economic burden of depression in 1990. *J Clin Psychiatry.* 1993;54:405–418.
59. Wickizer TM, Lessler D. Do treatment restrictions imposed by utilization management increase the likelihood of readmission for psychiatric patients? *Med Care.* 1998;36:844–850.

JOURNAL OF WOMEN'S HEALTH
Volume 13, Number 8, 2004

Effect of Patient Gender on Late-Life Depression Management

SUSAN M. FRAYNE, M.D., M.P.H.,[1] KATHERINE M. SKINNER, Ph.D.,[2,3,*] HAI LIN, M.D.,[3] ARLENE S. ASH, Ph.D.,[4] and KAREN M. FREUND, M.D., M.P.H.[4,5]

ABSTRACT

***Purpose*: To determine whether patient gender influences physicians' management of late-life major depression in older and younger elderly patients.**

***Methods*: In 1996–2001, physician subjects viewed a professionally produced videotape vignette portraying an elderly patient meeting diagnostic criteria for major depression, then answered interviewer-administered questions about differential diagnosis and treatment. Patient gender and other characteristics were systematically varied in different versions of the videotape, but clinical content was held constant. This was a stratified random sample of 243 internists and family physicians with Veterans Health Administration (VA) or non-VA ambulatory care practices in the Northeastern United States. Outcomes were whether physicians followed a guideline-recommended management approach: treating with antidepressants or mental health referral or both and seeing the patient for follow-up within 2 weeks.**

***Results*: Only 19% of physicians recommended treating depression (12% recommended antidepressants and 7% mental health referral), and 43% recommended follow-up within 2 weeks. Patient gender did not influence management recommendations in either younger old (67 year old) or older old (79 year old) patients ($p > 0.12$ for all comparisons).**

***Conclusions*: Gender disparities previously documented in the management of major conditions are not seen for the management of depression, a potentially stigmatized condition that does not require resource-intense interventions.**

INTRODUCTION

GENDER DISPARITIES IN INTENSITY of medical care have been scrutinized for over a decade.[1] Numerous studies find women less likely than men to receive intensive interventions for a range of conditions.[1–5] Much less is known about the effect of patient gender on the primary care management of major depression; indeed, the gender disparity effect could even be reversed for this specific condition. Given the high prevalence of depression women,[6,7] it is possible that under-

[1]Center for Health Care Evaluation, VA Palo Alto Healthcare System, and Stanford University, Palo Alto, California.
[2]School of Public Health, Boston University, Boston, Massachusetts.
[3]Center for Health Quality, Outcomes and Economic Research, VA Medical Center, Bedford, Massachusetts.
[4]Section of General Internal Medicine, and [5]Women's Health Unit, Evans Department of Medicine, Boston Medical Center, and Boston University, Boston, Massachusetts.
*Deceased.
This material is based on work supported by the Department of Veterans Affairs Health Services R&D grants GEN 97-023 and RCD 98-312, and National Institute on Aging grant AG12437-03. The views expressed in this paper are those of the authors and do not necessarily represent the views of the Department of Veterans Affairs.

recognition of depression[8] is less of a problem for female patients, who then have more of an opportunity than men to be treated.

To address this gap in the literature, we used an experimental design to determine whether physicians' management of late-life major depression varied as a function of patient gender. With women's longer life expectancy, aging is an important women's health issue,[9,10] and clinicians often undertreat depression in elderly patients.[11] Therefore, a secondary question was whether depression management varied by patient gender in younger old and older old age strata.

MATERIALS AND METHODS

Overview of experimental design

After viewing a professionally acted videotape vignette portraying an elderly patient who meets diagnostic criteria for major depression (in which patient gender and other characteristics were systematically varied but clinical content was held constant), practicing physicians answered structured, interviewer-administered questions about their diagnosis and management recommendations. This paper reports on one aspect of a larger study of late-life depression management in the medical setting, but the findings reported here about the effect of patient gender answer the primary question of the overall study. Study methods are described in detail elsewhere.[12,13]

Physician subjects

A random stratified sample of internists and family physicians with active ambulatory care practices were selected in New England (Connecticut, Rhode Island, Massachusetts, Vermont, New Hampshire, Maine); recruitment was extended to New York to assure an adequate sampling frame. Physicians were contacted by telephone to enlist their participation prior to a scheduled in-person interview. The Institutional Review Boards at New England Research Institute, Boston University, and each of the 10 participating Veterans Health Administration (VA) facilities approved this study, and physicians provided written informed consent.

In the first phase of our study, primary care physicians were randomly selected from Board of Registration in Medicine lists in 1996–1997 based on listed specialty (internal medicine or family practice). Sampling was stratified according to physician gender, specialty, years since medical school graduation, and race.

In a second recruitment phase, we replicated the study in a VA setting. The reason for replicating the study in VA was to examine whether findings would be consistent in a setting where women represent a minority population and where mental health issues are emphasized systemwide. This second phase of the study was initiated while recruitment for the non-VA phase was still ongoing. We randomly selected staff physicians from VA Department of Medicine/Ambulatory Care lists in 1999–2001, stratified according to physician gender and specialty status in lists provided by the Department of Medicine/Ambulatory Care at each VA facility (the size of the VA sampling frame did not make it possible to stratify on physician race). In VA hospitals, women are a minority population,[14] making VA a unique environment in which to examine gender differences in care. For the VA phase, we included not only primary care providers but also medical subspecialists because they are responsible for a substantial proportion of the primary care provided in VA.

Prospective subjects were recruited by telephone after receipt of an introductory letter asking them to participate in a study about "physicians' clinical decision making." Subjects who participated received $100 (or its equivalent in educational materials).

Interview procedures

Physician subjects viewed the depression vignette, then answered questions from a field-tested protocol (administered by a trained interviewer) about the probability of each condition in their differential diagnosis; what tests, referrals, or medications they would order; and when they would want to see the patient back for follow-up. The interview ended with items characterizing the physician's background and clinical practice. Physicians were not informed that the patient in the vignette was an actor and were unaware of the study hypotheses. The interviewer instructed the physician to use his or her own practice as the frame of reference when selecting management recommendations.

Videotaped vignettes

In a 5-minute, professionally produced videotape, an actor portrays a patient meeting *Diagnostic and Statistical Manual*, 4th ed (DSM-IV), criteria[15] for late-life major depression (anhedonia, anorexia, weight loss, insomnia, fatigue, psychomotor retardation, and thoughts of death). Because depression commonly occurs in the primary care setting with somatic rather than mood symptoms,[16] there was no mention in the vignette of feelings of depression or suicidal ideation. To maximize the realism of the vignette: (1) real primary care cases were used as the foundation for case development, and role-plays of these cases by clinicians were used to guide script development, (2) lay terms were used whenever possible, (3) three panels of practicing physicians were convened to confirm the authenticity of the script, and (4) an internist was present during filming to assure appropriate and consistent nonverbal cues. (The internist provided real-time feedback to actors on their nonverbal communication, requiring, for example, that they cast their eyes downward at the same point in the script.) In addition, we interviewed subjects in their clinical practice and asked them to manage the case as if the patient were coming to their own practice.

Patient gender was systematically varied in different versions of the videotape, as were other demographic characteristics (including age of the patient depicted, which was 67 or 79 years old, race of the patient, and in the non-VA group, socioeconomic status of the patient, indicated by occupation and clothing selection), using a previously developed methodology.[12,17] However, the clinical content and nonverbal communication were held constant.

Dependent variables

Physicians who included depression in their differential diagnosis and assigned it a probability $\geq 50\%$ were considered to have identified depression. Physicians were considered to have recommended guideline-based[18,19] treatment if they recommended an antidepressant or mental health referral or both at the initial visit. Monitoring was guideline-concordant if the physician recommended follow-up within 2 weeks. Of note, the Agency for Health Research and Quality (AHRQ) distributed its depression clinical guidelines in booklet format nationally in 1993, well before sample selection began.

Analysis

The primary, *a priori* objective of this study was to determine whether patient gender influenced the identification and management of depression. We completed bivariate analyses comparing rates of identification, treatment, and monitoring of depression for female vs. male patients depicted in the videotape and by patient gender-age dyads (older old female, older old male, younger old female, younger old male). We also conducted logistic regression analyses testing the effect of patient gender on the identification, treatment, and follow-up for depression after adjusting for physician characteristics (physician gender, years postgraduation from medical school, specialist status, VA vs. non-VA status). In sensitivity analyses, we repeated these analyses of the treatment and follow-up outcomes for the subset of physicians who reported that they provided primary care. Because baseline physician and caseload characteristics differed between the VA and non-VA physicians, we also conducted stratified analyses to evaluate for any confounding of findings by practice site (VA or non-VA). We used chi-square tests for discrete variables and t tests for continuous variables and conducted Fisher's exact test in instances of expected cell counts <5. The unit of analysis was the physician subject. We considered $p < 0.05$ as statistically significant.

For our main analyses (gender comparisons), we had $>80\%$ power to detect treatment differences of $\leq 15\%$. In analyses of interaction effects, our power to detect differences between groups was limited; these latter comparisons should be considered exploratory.

RESULTS

Among contacted physicians from our sampling frame, the response rate was 85% in phase 1 (128 non-VA physicians) and 78% in phase 2 (115 VA physicians), yielding a total sample of $n = 243$ physicians. The mean age of the full sample was 46.1 years, and 58% were male. The VA and non-VA physicians differed on practice characteristics: 50% vs. 26%, respectively, self-reported that they provided some specialty care; they saw a mean of 38 vs. 75 outpatients per week; 9% vs. 62% of their caseloads were female; and 69% vs. 26% of their caseloads were $>$ age 65 ($p < 0.001$ for these comparisons).

Table 1. Differences in Identification and Management of Depression, by Patient Gender in Vignette, for Full Cohort (n = 243)

	Patient gender in vignette		
Physician response	*Female* n = *120* *(%)*	*Male* n = *123* *(%)*	p
Identified depression	63.3	64.2	0.88
Treated depression	18.3	19.5	0.81
Antidepressant	11.7	13.0	0.75
Refer to mental health	6.7	8.1	0.66
2-Week follow-up	41.7	44.7	0.63

We found that 64% of physicians identified depression, defined as assigning at least a 50% probability to depression on their differential diagnosis. Based on the information in the vignette for an initial visit, 19% recommended treatment for depression, that is, 12% recommended antidepressant medication, and 7% recommended a mental health referral. Whereas standard clinical guidelines recommend follow-up within 2 weeks for patients with depression,[18,19] 43% of subjects recommended follow-up within this time frame (time to follow-up had a range of 2–150 days). Patient gender did not influence the likelihood that depression would be identified or that guideline-concordant treatment and follow-up would be recommended (Table 1). In logistic regression analyses adjusting for physician characteristics (physician gender, years postgraduation, specialty status, and VA or non-VA status), there continued to be no statistically significant gender effect: odds ratio (with 95% confidence interval) (OR[95% CI]) for identification of depression was 0.91 (0.53-1.55), for treatment was 0.93 (0.49-1.79), and for follow-up within 2 weeks was 0.89 (0.51-1.46).

There were also no systematic differences in identification, treatment, or follow-up of depression by gender-age dyad. Differences noted were not statistically significant and were not consistent across age group or gender to suggest a pattern of management differences based on age or gender (Table 2).

Among the subset who provided primary care, there were likewise no gender differences in treatment or follow-up. Despite differences in caseload between VA vs. non-VA physicians and despite the fact that there were isolated differences in depression management among VA vs. non-VA physicians (reported elsewhere[13]), the lack of effect of patient gender on depression identification and management was comparable in VA and non-VA settings.

DISCUSSION

Many physicians recognized late-life depression in a realistic videotaped clinical scenario of an elderly patient with somatic symptoms who met the criteria for the diagnosis. However, despite the availability of national depression management guidelines that were widely disseminated at least 3 years prior to our first interview,[18,19] most subjects did not recommend early intervention or early follow-up. This lack of recommended intensive management is independent of patient gender, and exploratory analyses do not detect a gender effect in either older or younger patients with late-life depression.

We found a lack of effect of patient gender on depression management. This is consistent with prior work indicating a lack of effect of patient gender on depression treatment (with pharmacotherapy or psychotherapy) by psychiatrists, al-

Table 2. Differences in Identification and Management of Depression, by Patient Age-Gender Dyad, for Full Cohort (n = 243)

	79-year-old patient		*67-year-old patient*		
Physician response	*Female* n = *61* *(%)*	*Male* n = *59* *(%)*	*Female* n = *59* *(%)*	*Male* n = *64* *(%)*	p[a]
Identified depression	65.6	62.7	61.0	65.6	0.94
Treated depression	26.2	17.0	10.2	21.9	0.13
Antidepressant	16.4	10.2	6.8	15.6	0.32
Refer to mental health	9.8	8.5	3.4	7.8	0.56
2-Week follow-up	36.1	44.1	47.5	45.3	0.61

[a]p value for the chi-square test comparing the four age-gender groups.

though that study was limited by a low response rate.[20] In the primary care setting, this issue has received little prior attention, although one study suggested that women were more likely than men to receive antidepressant and anxiolytic prescriptions for symptoms of depression in primary care.[21] Although that observational study used the available information to control for patient presentation and address the elevated base rates of depression in women as carefully as possible, the author acknowledged the need for future experimental studies to determine if his findings reflected physician bias vs. unmeasured confounders.[21] Our experimental design addressed this need.

Of note, examining the effect of patient gender on depression management was the primary objective of our overall study, and our study was designed to have ample statistical power to detect important gender differences. Therefore, although we observed no difference in the management of depression for female vs. male patients, we consider it important to report this negative finding so as to avoid contributing to publication bias.[22,23]

Interestingly, our findings contrast with a consistent pattern in many other studies suggesting that women are less likely than men to receive intensive interventions for a range of conditions.[1–4] Our study was not designed to understand the reasons for lack of gender differences, but our findings raise intriguing possibilities. First, it is possible that gender disparities in intensity of treatment are most pronounced for resource-intensive interventions, such as invasive cardiovascular procedures, kidney transplants, colonoscopies, or state-of-the-art HIV treatments.[1–5] Indeed, for such simple interventions as frequency of visits or routine laboratory testing, women tend to receive more care than men.[1,24,25] Our finding may reflect the fact that depression management requires lower intensity interventions. Second, some studies find that clinicians are more likely to attribute women's physical symptoms to psychological causes, sometimes overdiagnosing psychiatric illness in women.[26,27] Thus, although we found no effect of patient gender on the identification of depression, physicians might have a heightened responsiveness to depression in women. This could offset any general tendency to offer less interventions to women.

The overall low rate of intervention that we report is consistent with prior work documenting low treatment rates for depression in elderly patients despite the morbidity of untreated disease and the availability of effective treatments.[11,28,29] There are three major potential explanations for this observation. The first relates to patient preferences, that is, the possibility that elderly patients decline recommended treatment for depression. Our experimental design eliminated this explanation, as physicians' recommendations were independent of patient feedback. The second potential explanation is that the higher prevalence of chronic illness, physical disability, and loss of independence seen with advancing age adds complexity to depression management and promotes therapeutic nihilism.[30,31] In our study, this would not be expected to contribute, as the patients depicted in the videotape presented identically: for all gender-age dyads, the patient was married, living independently, and without serious chronic illness. The third possible explanation is that there is a direct effect of ageism, with elderly patients being selectively undertreated. Our study design was ideally suited to examine this possibility. Although we documented low rates of recommending treatment, we found no age gradient. Treatment recommendations were similar whether the physician was assigned to a patient age 67 or 79 years old. This does not exclude the possibility that there is a threshold effect of age on intensity of treatment, with the age of patients depicted in the vignette all falling above this threshold.

Our methodology has several particular strengths. We used a rigorous experimental approach that overcomes case mix problems[32] inherent in the observational designs of earlier work.[20,21] This allowed us to directly compare the management of female and male older old and younger old patients with identical symptoms. Sample size was chosen to assure that there would be ample statistical power to detect any meaningful differences in the management of female and male patients and older old and younger old patients. The methodology used in this study has been applied to a variety of questions and has been shown in work by us and others to be capable of detecting important differences in clinical management in different patient groups and in different physician groups.[12,13,17,33] Finally, by examining the management of older old patients with depression, we address an age group for whom the research literature is particularly sparse.[34]

Limitations

Several limitations of this approach should also be raised. The first relates to external validity: physicians' actual practice might differ from their responses to the videotaped scenario. For example, the fact that the patient did not report suicidal ideation or that physicians did not have the opportunity to personally interview the patient in the vignette (e.g., to obtain additional information about the severity of the depression) could have led some to hesitate to initiate treatment. However, clinical vignettes have good validity when compared with other methodologies for measuring quality of care.[35] We took many steps to assure the vignettes' realism and to facilitate physicians' viewing the case as a patient in their own practices. Indeed, physicians made spontaneous comments: "I saw that patient this morning." Second, it is important to recognize that physicians' responses reflected their recommendations for a single visit with a continuity patient. It is possible that they would have initiated treatment at follow-up visits. However, early intervention is the standard of care for depression,[18,19] and less than half of the subjects planned to see the patient in follow-up within 2 weeks. Third, the patient depicted in the video was living independently and was relatively healthy, and all physician subjects practiced in the Northeastern United States. The generalizability of our findings to institutionalized or infirm elderly patients and to physicians practicing in other geographic regions is unknown.

CONCLUSIONS

Our study points to a continuing need for more aggressive treatment of depression in all elderly patients, independent of gender, supporting the importance of recent studies examining promising interventions for depression in the primary care setting.[36–40] Unlike the gender effects reported previously for problems requiring intensive intervention, we found that patient gender and patient gender-age dyad do not influence management of late-life depression, a potentially stigmatized condition for which well-established, low-intensity treatments are available. These findings provide additional clues in the disparities literature and can serve as a foundation for future work examining why such disparities are seen for some but not all conditions.

ACKNOWLEDGMENTS

John B. McKinlay, Ph.D., Julie T. Irish, Ph.D., and Ting H. Lin, Ph.D., of New England Research Institute were central to the design and execution of the first phase of this study. We are grateful for the extensive technical assistance provided by Stephanie J. Smith, Laurie Craigen, Jeff Murphy, Patricia Loya, and Danielle Pierce. We also appreciate the insights of Mary Goldstein, M.D., M.S., and Rachel Kimerling, Ph.D., who reviewed earlier versions of this paper.

REFERENCES

1. Council on Ethical and Judicial Affairs, American Medical Association. Gender disparities in clinical decision making. JAMA 1991;266:559.
2. Ayanian JZ, Epstein AM. Differences in the use of procedures between women and men hospitalized for coronary heart disease. N Engl J Med 1991;325:221.
3. Ramani S, Byrne-Logan S, Freund KM, Ash A, Yu W, Moskowitz MA. Gender differences in the treatment of cerebrovascular disease. J Am Geriatr Soc 2000; 48:741.
4. McMahon LF Jr, Wolfe RA, Huang S, Tedeschi P, Manning W Jr, Edlund MJ. Racial and gender variation in use of diagnostic colonic procedures in the Michigan Medicare population. Med Care 1999;37: 712.
5. Andersen R, Bozzette S, Shapiro M, et al. Access of vulnerable groups to antiretroviral therapy among persons in care for HIV disease in the United States. HCSUS Consortium. HIV Cost and Services Utilization Study. Health Serv Res 2000;35:389.
6. 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.
7. Weissman MM, Leaf PJ, Tischler GL, et al. Affective disorders in five United States communities. Psychol Med 1988;18:141.
8. Perez-Stable EJ, Miranda J, Munoz RF, Ying YW. Depression in medical outpatients. Underrecognition and misdiagnosis. Arch Intern Med 1990;150:1083.
9. La Croix AZ, Newton KM, Leveille SG, Wallace J. Healthy aging. A women's issue. West J Med 1997; 167:220.
10. Task Force on Older Women's Health. American Geriatrics Society. Older women's health. J Am Geriatr Soc 1993;41:680.

11. Lebowitz BD, Pearson JL, Schneider LS, et al. Diagnosis and treatment of depression in late life. Consensus statement update. JAMA 1997;278:1186.
12. Feldman HA, McKinlay JB, Potter DA, et al. Nonmedical influences on medical decision making: An experimental technique using videotapes, factorial design, and survey sampling. Health Serv Res 1997; 32:343.
13. Frayne S, Freund K, Skinner K, Ash A, Moskowitz M. Depression management in medical clinics: Does healthcare sector make a difference? Am J Med Qual 2004;19:28.
14. Weiss TW, Ashton CM. Access of women veterans to Veterans Affairs hospitals. Women Health 1994;21:23.
15. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, 4th ed. Washington, DC: American Psychiatric Press, 1994.
16. Levenson J, ed. Depression. Philadelphia: American College of Physicians, 2000.
17. McKinlay JB, Lin T, Freund K, Moskowitz M. The unexpected influence of physician attributes on clinical decisions: Results of an experiment. J Health Soc Behav 2002;43:92.
18. Depression-Guideline Panel. Depression in primary care: Volume 1. Detection and diagnosis. Clinical practice guideline number 5. Rockville, MD: U.S. Department of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research, April 1993. AHCPR Publication Number 93-0550.
19. Depression-Guideline Panel. Depression in primary care: Volume 2. Treatment of major depression. Clinical practice guideline, Number 5. Rockville, MD: U.S. Department of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research, April 1993. AHCPR Publication Number 93-0551.
20. Olfson M, Zarin DA, Mittman BS, McIntyre JS. Is gender a factor in psychiatrists' evaluation and treatment of patients with major depression? J Affect Disord 2001;63:149.
21. Hohmann AA. Gender bias in psychotropic drug prescribing in primary care. Med Care 1989;27:478.
22. Dickersin K, Min YI, Meinert CL. Factors influencing publication of research results. Follow-up of applications submitted to two institutional review boards. JAMA 1992;267:374.
23. Rennie D. Fourth International Congress on Peer Review in Biomedical Publication. JAMA 2002;287:2759.
24. Verbrugge LM, Steiner RP. Physician treatment of men and women patients: Sex bias or appropriate care? Med Care 1981;19:609.
25. Meeuwesen L, Schaap C, van der Staak C. Verbal analysis of doctor-patient communication. Soc Sci Med 1991;32:1143.
26. Bernstein B, Kane R. Physicians' attitudes toward female patients. Med Care 1981;19:600.
27. Redman S, Webb GR, Hennrikus DJ, Gordon JJ, Sanson-Fisher RW. The effects of gender on diagnosis of psychological disturbance. J Behav Med 1991;14: 527.
28. Blazer DG. Depression in the elderly. Myths and misconceptions. Psychiatr Clin North Am 1997;20:111.
29. Jenike MA. Psychiatric illnesses in the elderly: A review. J Geriatr Psychiatry Neurol 1996;9:57.
30. Lyness JM, Bruce ML, Koenig HG, et al. Depression and medical illness in late life: Report of a symposium. J Am Geriatr Soc 1996;44:198.
31. Montano CB. Primary care issues related to the treatment of depression in elderly patients. J Clin Psychiatry 1999;60(Suppl 20):45.
32. Iezzoni L. Risk adjustment for measuring healthcare outcomes, 2nd ed. Chicago: Health Administration Press, 1997.
33. Freund KM, Moskowitz MA, Lin TH, McKinlay JB. Early antidepressant therapy for elderly patients. Am J Med 2003;114:15.
34. Bugeja G, Kumar A, Banerjee AK. Exclusion of elderly people from clinical research: A descriptive study of published reports. Br Med J 1997;315:1059.
35. Peabody JW, Luck J, Glassman P, Dresselhaus TR, Lee M. Comparison of vignettes, standardized patients, and chart abstraction: A prospective validation study of 3 methods for measuring quality. JAMA 2000; 283:1715.
36. 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.
37. Wells KB, Sherbourne C, Schoenbaum M, et al. Impact of disseminating quality improvement programs for depression in managed primary care: A randomized controlled trial. JAMA 2000;283:212.
38. 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.
39. van Os TW, Ormel J, van den Brink RH, et al. Training primary care physicians improves the management of depression. Gen Hosp Psychiatry 1999;21:168.
40. Meredith LS, Jackson-Triche M, Duan N, Rubenstein LV, Camp P, Wells KB. Quality improvement for depression enhances long-term treatment knowledge for primary care clinicians. J Gen Intern Med 2000; 15:868.

Address reprint requests to:
Susan Frayne, M.D., M.P.H.
Center for Health Care Evaluation
795 Willow Road (152-MPD)
Menlo Park CA 94025

E-mail: sfrayne@stanford.edu

The Views of U.S. Medical School Deans Toward Academic Primary Care

Robert H. Friedman, MD, Sandhya Wahi-Gururaj, MD, MPH, Joel Alpert, MD, Howard Bauchner, MD, Larry Culpepper, MD, MPH, Timothy Heeren, PhD, and Allen Singer, PhD

ABSTRACT

Purpose. To understand the views of U.S. medical school deans about their primary care faculties.

Method. In 2000, the authors mailed a questionnaire containing 43 multipart items to deans of 130 U.S. allopathic medical schools. The questionnaire assessed the deans' attitudes about and evaluation of primary care at their school and their school's efforts to strengthen it. Deans were asked to compare family medicine, general internal medicine, and general pediatrics with nonprimary care clinical departments at their schools.

Results. Of the 83 (64%) deans who responded, 82% reported their school had departments or divisions of family medicine, general internal medicine, and general pediatrics. Deans rated general internal medicine and general pediatrics higher than nonprimary care faculty on clinical expertise and productivity ($p < .001$) and family medicine equivalent to nonprimary care faculty. Deans rated all three primary care faculties superior to nonprimary care faculty for teaching skills ($p < .001$) and programs ($p < .05$), but lower than nonprimary care disciplines for research productivity ($p < .01$) and revenues ($p < .001$). They rated family medicine and general pediatrics lower for research skills ($p < .001$), but 73% of deans stated research was equally important for primary care and nonprimary care departments. Deans considered overall financial resources to be equivalent for primary care and nonprimary care departments, but 77% of deans felt primary care departments or divisions needed financial support from the medical school to survive. Most deans attempted to strengthen primary care by changing the curriculum to promote primary care and by providing financial support.

Conclusions. Deans ranked primary care faculty high on clinical and teaching measures. Although they considered research to be an important activity for primary care faculty, they evaluated it low relative to nonprimary care departments.

Acad Med. 2004;79:1095–1102.

The role of primary care physicians (family medicine, general internal medicine, and general pediatrics) in academic health centers has been well established. Primary care faculty have a broad scope of responsibilities in patient care, teaching, research, and administration.[1–3] Generalist–physician faculty serve as role models for medical students and have major responsibility for the education of future primary care clinicians.[4,5] Medical students perceive higher levels of encouragement to enter primary care careers at schools with strong primary care orientations.[6] The training of future generalists has received attention in the past three decades because of a reported shortage of primary care clinicians in the physician workforce.[7–10] Although the primary care workforce has grown since 1970, in 1999 the Council on Graduate Medical Education reported there is still a need to produce more generalists.[11]

Previously identified challenges to academic generalism include its financial vulnerability, weak institutional influence and prestige, limited role in medical education, and inadequate scholarly output.[3] Knowing whether these factors persist is important because they potentially threaten the ability of primary care de-

Dr. Friedman is professor and Dr. Wahi-Gururaj is assistant professor, Department of Medicine; Dr. Alpert is professor and Dr. Bauchner is professor, Department of Pediatrics; Dr. Culpepper is professor and chair, Department of Family Medicine; all at Boston University School of Medicine, Boston, Massachusetts. Dr. Heeren is professor, Department of Biostatistics, Boston University School of Public Health; and Dr. Singer is vice president, Research and Information Services, American Association of Colleges of Osteopathic Medicine, Chevy Chase, Maryland.

Correspondence should be addressed to Dr. Friedman, Boston Medical Center, Boston University School of Medicine, Medicine Information Systems Unit, 720 Harrison Avenue, Suite 1102, Boston, MA 02118; e-mail: (rfriedma@bu.edu).

partments and divisions to function adequately in education, scholarship, and health services delivery.

Determining the medical school deans' attitudes toward and evaluation of primary care in their institutions and in medical schools nationally is critical to our understanding the issues facing primary care medicine in academia. In 1999, deans of half of the U.S. allopathic medical schools reported it was more important for their institutions to train future primary care physicians than to train future specialist physicians.[12] We do not know whether this supportive view of primary care translates into supportive medical school policies and structures and into strong primary care faculties and programs. In this study, we surveyed the deans of all U.S. allopathic medical schools to better understand these matters.

Method

Design

In 2000, we mailed questionnaires to deans of the 130 allopathic medical school campuses in the United States and the Commonwealth of Puerto Rico. To identify the names, addresses, and telephone numbers of the deans, we accessed medical school Web sites and supplemented this search with telephone inquiries to the schools. Nonrespondents were contacted by telephone two weeks after the mailing and replacement questionnaires were sent if necessary. The questionnaire contained 43 multipart items regarding the organization of primary care at the school and in departments and divisions of family medicine, general internal medicine, and general pediatrics; the dean's attitudes toward academic primary care; the dean's evaluation of the primary care faculty, programs, and resources; and the school's efforts to strengthen academic primary care at their institution. The instrument was pilot tested by six former allopathic medical school deans. The Boston University School of Medicine Institutional Review Board approved the study.

Data and Variables

To understand the organization of academic primary care at each school, we inquired whether the school had a formal department of family medicine and divisions of general internal medicine and general pediatrics or their equivalents. We also asked about the nature of the current relationships among these departments and divisions. We inquired whether a dean for primary care existed at the school, collected information on the current number of full-time doctoral faculty, and asked the deans to project growth over the next five years in both the primary care departments and divisions and in other clinical and nonclinical departments.

We inquired about the percentages of the preclinical curriculum (traditional medical school years one and two) and the clinical curriculum (traditional medical school years three and four) devoted to primary care and the proportion taught by primary care faculty. In addition, we asked about the degree of coordination or integration of the generalist department and divisions in teaching these curricula.

To understand the deans' attitudes, we asked them the degree to which they agreed or disagreed with 11 indicator statements addressing finances, research, and the role and influence of primary care faculty at the medical school.

The deans' evaluated each of the primary care disciplines (family medicine, general internal medicine, and general pediatrics) at his or her institution in relation to the norm for the school's nonprimary care clinical departments. They were asked to consider 12 factors pertaining to the clinical, educational, and research roles of the generalist faculty and the financial resources of each primary care academic unit. The factors were assessed on six-point Likert-type scales (1 = "worse than nonprimary care clinical departments" to 6 = "better than nonprimary care clinical departments"). We compared the deans' mean ratings to a score of 3.5 (the midpoint on the scale) to determine whether the particular primary care department or division was viewed more or less favorably than other clinical departments. Thus, a mean score greater than 3.5 would indicate that deans rated the generalist department or division more favorably than other clinical departments, whereas a score less than 3.5 would indicate a less favorable rating. In addition, we compared the deans' ratings across family medicine, general internal medicine, and general pediatrics. We combined the 12 scores for each factor, giving equal weights to each factor, to create a summative measure.

Finally, we assessed the measures the medical school had taken to strengthen its primary care departments, divisions, programs, and faculty. We inquired about 13 potential interventions in medical school policy, medical education, and administrative and financial support.

To assess the representativeness of our sample, we compared the characteristics of the medical schools of responding and nonresponding deans using data provided by the Association of American Medical Colleges (AAMC).

We performed sensitivity analyses by determining what impact certain medical school characteristics would have on the deans' responses (e.g., public versus private, the medical schools' research intensity). An institution was defined as "research intensive" if greater than or equal to 30% of the medical school's revenues were from research grant dollars as reported in the AAMC's database, an approximate median split.

Analyses

Responses are described through means and standard deviations for contin-

Table 1

Characteristics of the Organization of Primary Care within 83 U.S. Allopathic Medical Schools Based on the Responses of Medical School Deans, 2000

Characteristic	Response	
Formally organized department or division (%)		
Family medicine	92	
General internal medicine	90	
General pediatrics	86	
Institutions with family medicine department and general internal medicine and general pediatrics divisions	82	
Full-time faculty, mean (SD)		
Total faculty (no. = 73)	653 (474)	
Faculty in clinical departments (no. = 71)	563 (456)	
Primary care faculty, mean (% of clinical faculty)		
Family medicine	28 (9)	
General internal medicine	55 (11)	
General pediatrics	36 (8)	
Combined family medicine, general internal medicine, general pediatrics	124 (28)	
Inter-relationships among the three primary care disciplines, no. (%)	Current	Preferred
Wholly independent	17 (21)	2 (3)
Some coordination	39 (48)	25 (31)
Some integration	24 (30)	49 (61)
Complete integration	1 (1)	5 (6)

uous and ordinal-scaled variables and through numbers and percentages for categorical variables. For some continuous and ordinal-scaled variables, we categorized the responses. Using the finite population correction factor, based on overall sampling from 64% schools, 95% confidence intervals for the percentages reported are accurate to within 3%.

The deans' ratings of family medicine, general internal medicine, and general pediatrics relative to nonprimary care clinical departments at the medical school in general were compared through a one-factor repeated measures analysis of variance. We used a two-way repeated measures analysis of variance to determine the impact of medical school characteristics. Where overall differences in ratings were significant (at $p < .05$), pairwise comparisons were performed using Tukey's procedure to identify specific differences in ratings.

Results

Medical Schools' Characteristics

Deans from 83 (64%) allopathic medical schools responded, and at least 76 (58%) responded to each question, unless otherwise indicated.

Fifty-two percent of the deans had assumed their positions since 1996, 25% since 1999. Most deans had professional backgrounds in clinical departments; the largest percentages were from departments of medicine (40%) or psychiatry (10%), and fewer were from departments of pediatrics or family medicine (8% and 4%, respectively). Eight percent of the deans previously belonged to a basic medical science department.

Sixty-one percent of the responding deans were from private medical schools. Thirty-five percent of the medical schools were located in the South, 23% in the Northeast, 26% in the Midwest, and 16% in the West. The schools had a median of 473 medical students, 635 full-time faculty, and total revenues of $128 million. There were no statistically significant differences between responding and nonresponding medical schools regarding these variables.

Organization of Primary Care

Table 1 shows the proportion of schools with departments of family medicine, divisions of general internal medicine, and general pediatrics, their interrelationships, and the numbers and percentages of faculty in these primary care departments and divisions.

Eighty-three percent of the deans anticipated an increase in the number of full-time doctoral level faculty in the primary care departments or divisions over the upcoming five years, which was similar to the deans' estimates for all clinical departments (82%) and nonclinical departments (82%) at their medical schools over the same period (data not shown). The mean projected change in number of faculty was 12%. There were no differences in projected changes for clinical versus nonclinical department faculty; for primary care versus nonprimary care faculty; or between faculty in family medicine, general internal medicine, and general pediatrics.

The deans' characterizations of the interrelationships among the three generalist departments and divisions are also shown in Table 1. Only one medical school had a fully integrated primary care department. Deans indicated they preferred more integration of primary care activities than their institutions had at present. Sixty-one percent preferred some integration, but only five (6%) preferred complete integration.

A dean for primary care existed in 14 (17%) medical schools. This dean provided leadership for the school's primary care curriculum for medical students (12 schools), for primary care research (ten schools), and for relationships with pri-

Table 2

Percentages of Agreement of 83 U.S. Allopathic Medical School Deans with Indicators Regarding Academic Primary Care, 2000*

Indicator Statement	% Agree
It is more difficult for faculty without "bench" science research backgrounds to be effective medical school leaders	22
The main role of primary care is to bring in patients that support the institution's clinical and educational enterprises	30
There is a great deal of conflict and competition among the primary care disciplines in medical schools	34
Generalists should be the core faculty of the medical school	46
Primary care faculty and their departments/divisions exert insufficient influence on the policies and priorities of most medical schools	49
There are not enough primary care role models for medical students to emulate	60
In general, primary care faculty are not held in high esteem by other faculty	60†‡
It is appropriate for government to preferentially support primary care faculty and education	65†‡
Primary care needs financial support from the medical school to survive	77†
There is a rich research and intellectual base for academic primary care	63
Research is as important for primary care departments/divisions as it is for other departments/divisions	73†

*Responses on a six-point Likert-type scale, grouped into two categories: 1–3 = disagree and 4–6 = agree.

†Deans of public institutions > deans of private institutions ($p < .05$).

‡Deans of non–research-intensive institutions > deans of research-intensive institutions ($p < .05$).

mary care physicians and services in the community (ten schools). At seven of these schools, this dean was responsible for the primary care curriculum for residents, and at five schools the dean was responsible for primary care clinical services. Seven of these schools provided a separate budget for the primary care dean.

Fifteen percent of the preclinical curriculum, on average, was devoted to primary care (range 0–90%), and primary care faculty taught 25% of this curriculum (range 0–100%). Approximately 60% of deans reported moderate to high levels of integration among their primary care departments and divisions for teaching the preclinical curriculum; 13% had no coordination. On average, a higher proportion of the clinical curriculum was devoted to primary care (37%, range 7–70%) and taught by primary care faculty (39%, range 8–100%). The level of integration among generalist departments and divisions in teaching this clinical curriculum was moderate to high in 34% of institutions (whereas 17% reported no coordination).

Deans' Attitudes toward Academic Primary Care

The deans' responses to 11 indicator statements on academic primary care are shown in Table 2. Only 30% agreed with the view that "the main role of primary care departments/divisions is to bring in patients that support the institution's clinical and educational enterprise." Forty-six percent thought that "generalists should be the core faculty of the medical school," but 60% believed that there were not enough generalist faculty role models. Deans of public and non–research-intensive institutions were more likely to agree that primary care faculty are not held in high esteem ($p = .01$ and .02, respectively). There were no other significant differences in attitudes by medical school characteristics.

When probed about financial issues, 65% of deans agreed with the viewpoint that "it is appropriate for state and federal governments to preferentially support primary care faculty and primary care education," while 77% agreed that primary care departments and divisions need financial support from the medical school to survive. Deans of public and non–research-intensive institutions were more likely to agree with the former statement ($p < .01$, $p = .02$, respectively), while deans of public medical schools were more likely to agree with the latter one ($p < .01$). Deans of public institutions were more likely to agree that research was equally important for primary care departments and divisions ($p < .01$). Otherwise, there were no significant differences in level of agreement by medical school characteristics.

Deans' Evaluation of the Primary Care Disciplines

The deans' evaluations of primary care faculty, programs, and financial resources are shown in Table 3. They rated faculty in divisions of general pediatrics and general internal medicine higher than, and family medicine faculty equal to, nonprimary care clinical faculty on their clinical expertise and productivity. They judged the faculty's teaching skills in all three primary care disciplines superior to those of nonprimary care clinical faculty. Faculty in general pediatrics and general internal medicine were rated higher than specialists on educational productivity, whereas family medicine faculty were rated equal to. The deans viewed the research skills and productivity less favorably for general pediatrics and family medicine faculty compared with nongeneralists. Faculty in general internal medicine were considered equivalent in their research skills, but were rated lower for their faculty's research productivity.

Table 3

Mean (SD) Responses of 83 U.S. Allopathic Medical School Deans to Questions about Primary Care Disciplines' Faculty, Programs, and Financial Resources Compared with Nonprimary Care Clinical Departments, 2000*

Characteristic	Family Medicine† Mean (SD)	General Pediatrics†‡ Mean (SD)	General Internal Medicine† Mean (SD)	*p* Value‡§
Faculty clinical expertise	3.4 (.8)	4.0 (.7)[z]	4.3 (.8)[z]	.001[a,b,c]
Faculty clinical productivity	3.3 (1.0)	3.8 (.9)[y]	3.9 (.9)[z]	.001[a,b]
Faculty teaching skills	4.3 (.9)[z]	4.4 (.8)[z]	4.6 (.9)[z]	.107
Faculty educational productivity	3.7 (.9)	3.9 (.9)[z]	4.1 (1.0)[z]	.002[b]
Faculty research skills	2.5 (1.0)[z]	3.0 (1.2)[z]	3.3 (1.1)	.001[a,b]
Faculty research productivity	2.4 (1.0)[z]	2.9 (1.1)[z]	3.1 (1.2)[y]	.001[a,b]
Quality of medical student teaching	4.4 (.9)[z]	4.5 (.9)[z]	4.6 (.8)[z]	.207
Quality of internship/residency program¶	3.8 (1.2)[x]	4.5 (.9)[z]	4.3 (1.0)[z]	.001[a,b]
Quality of fellowship program¶	3.0 (1.1)[y]	3.8 (1.1)	3.9 (.9)[z]	.001[a,b]
Level of clinical revenues	2.7 (1.0)[z]	3.1 (1.1)[y]	3.4 (1.0)	.001[a,b]
Level of research revenues	2.2 (1.0)[z]	2.6 (1.1)[z]	2.9 (1.2)[z]	.001[a,b,c]
Overall financial resources	3.4 (1.1)	3.4 (1.0)	3.5 (1.0)	.571
Summative measure	3.3 (.7)[y]	3.7 (.7)[x]	3.9 (.2)[z]	< .001[a,b]

*Six-point Likert-type scale (1 = worse than nonprimary care clinical departments, 6 = better than nonprimary care clinical departments). A score of <3.5 or >3.5 (midpoint of scale) indicates worse than or better than nonprimary care disciplines, respectively.

†Comparison with the average nonprimary care clinical departments significant at x = $p < .05$, y = $p < .01$, z = $p < .001$.

‡The response rate for general pediatrics ≥56% (n = 73–78), except for response on fellowship programs.

§Tukey's multiple comparisons at an experiment-wise alpha .05 show significant differences between means for: a = family medicine and general pediatrics, b = family medicine and general internal medicine, and c = general pediatrics and general internal medicine.

¶Research-intensive institutions more likely to rate family medicine quality lower, $p < .01$ for both.

Among the educational programs, medical student teaching and residency programs received superior ratings for each primary care discipline. The ratings for the fellowship programs varied, with family medicine fellowships viewed less positively than, general pediatrics fellowships equal to, and general internal medicine more favorably than the norm for the nonprimary care clinical disciplines.

All three generalist disciplines were rated lower than were the nonprimary care disciplines with respect to their level of research revenues. Deans judged clinical revenues to be lower for family medicine and general pediatrics, and equivalent for general internal medicine, compared with the other clinical disciplines. However, the deans considered the overall financial resources (which included institutional support) to be equivalent for generalist and nongeneralist departments and divisions alike.

Deans rated general internal medicine and general pediatrics higher than family medicine on a summative measure for all 12 areas ($p < .001$). They ranked general internal medicine significantly higher than family medicine in nine of 12 areas, rating the two disciplines to be equivalent for faculty teaching skills, quality of medical student teaching, and overall financial resources. The deans rated general pediatrics higher than family medicine for eight of 12 areas, rating faculty educational skills and productivity, the quality of medical student education, and overall financial resources equally. They rated general internal medicine higher than general pediatrics in two areas: faculty clinical expertise and level of research revenues.

Public versus private status did not play a significant role in the deans' evaluations of the primary care disciplines. Deans of research-intensive institutions rated the quality of family medicine residency training lower (mean 3.3 versus 4.1) and fellowships lower (mean 2.5 versus 3.6) than did deans of non–research-intensive institutions ($p < .01$ for both).

Measures to Strengthen Primary Care Departments and Divisions

Most deans reported their schools attempted to strengthen primary care departments and divisions by changing the curriculum to promote primary care. To this end, they established required third-year clerkships in primary care, increased medical students' training in ambulatory care settings, and added curriculum time for primary care in the undergraduate years (see Table 4). Three-quarters of medical schools placed a major emphasis on financially supporting primary care departments and divisions and faculty as a means to strengthen primary care at their institu-

Table 4

Percentages of 83 Allopathic Medical School Deans Reporting Measures to Strengthen Primary Care Departments, 2000

Characteristic	% Placing Major Emphasis*
Interventions on curriculum	
Required third-year primary care clerkship	90†
Increasing medical student training in ambulatory setting	89
Increasing primary care curriculum time in undergraduate medical education	71
Providing resources	
Financially supporting primary care	76
Increasing primary care faculty	65
Administrative interventions	
Changing promotion/tenure policy to recognize nonresearch performance	64
Increasing primary care faculty on admissions committees	60
Increasing primary care faculty influence	60
Upgrading administrative status of primary care	39
Strengthening primary care administrative structure in dean's office	38
Other actions	
Promoting health services and clinical epidemiology research	65
Facilitating cooperation among primary care departments/divisions	56
Primary care orientation a consideration for medical school admission	40

*Seven-point Likert-type scale responses grouped as 0 = no emphasis, 1–3 = minor emphasis, and 4–6 = major emphasis.

†More public schools required third-year primary care clerkships ($p < .01$).

tions. Fewer focused on administrative or policy changes such as upgrading the administrative status of primary care departments and divisions, strengthening primary care administrative structure at the dean's level (25% of schools placed no emphasis on this method), and facilitating cooperation among primary care departments and divisions. There were no major differences by medical schools' characteristics.

Discussion

In their responses in our national study, medical school deans had a moderately positive assessment of their primary care departments and divisions, but they identified weaknesses that ought to be of concern to medical educators, policymakers, and the general public. Our results indicate that two of the weaknesses previously identified in academic generalism (i.e., inadequate scholarly output and financial vulnerability[3]) continue to exist. Furthermore, as reported in the past, deans had mixed views regarding the level of institutional influence and prestige of academic primary care departments and divisions. The limited role of primary care faculty in medical education that had been previously reported appears no longer to be the case.

Limited research has been a problem for academic primary care departments and divisions since their establishment in the 1960s and 1970s.[3] That this problem persists is evident in the deans' negative evaluations of primary care faculty's research skills and productivity as well as the level of research revenues of generalist units. Yet, most deans in our study stated that research was important for primary care faculty. Several factors, including too few qualified applicants for physician–investigator positions and insufficient faculty time and financial support for research, have been shown to negatively affect research in primary care[3,13–15] and the health of clinical research in general.[16] Because inadequate scholarship, whether real or perceived, continues to plague academic primary care, medical school and teaching hospital leaders should undertake efforts to further understand the causes and institute remedies.

Although deans viewed primary care departments and divisions to be financially vulnerable in their ability to generate income, they believed their overall financial resources were equivalent to other clinical departments. This parity appears to be achieved by a combination of institutional financial support and preferential policies by governmental agencies, strategies with which the majority of deans agreed. The impact of governmental support has been noteworthy. Title VII funding of the Public Health Service Act has been associated with the growth of the family medicine physician workforce as well as higher rates of practice in underserved areas,[17,18] and 86 family medicine programs received Title VII funding in 2002.[19] An ongoing question is whether this support will continue given concerns raised over potential cuts in Title VII funding.

Negative attitudes towards primary care in academic health centers have been described previously.[20] The level of institutional influence and prestige of generalist departments or divisions in our study was mixed, at best, with deans evenly divided on whether primary care faculty and departments or divisions exerted sufficient influence on medical schools' policies. Many deans indicated the need for additional primary care role models and recognized that generalist faculty were not held in high esteem by other medical school faculty.

The role of primary care faculty in medical education appears to be less of an issue than in the past, as we found one-quarter of the preclinical and more than one-third of the clinical curricu-

lum, on average, being taught by primary care faculty. Deans reported very positive assessments of the educational skills, productivity, and programs of the generalist faculty. The deans' positive assessment of role of generalist faculty in teaching students is encouraging given their desire to emphasize primary care in medical school curricula[21] and the fact that increased exposure to generalist faculty has been associated with students choosing primary care careers.[4,5]

Our findings regarding the deans' evaluation of and attitudes towards academic primary care faculty are generally concordant with those reported by Block et al. in 1996.[20] In their study of primary care in academia, 90% of deans rated primary care teaching as good as or better than teaching in other disciplines, but only 37% considered primary care research to be as good as or better than research performed in other disciplines.

Most medical schools have formally organized generalist departments or divisions. The presence of a department of family medicine has been shown to positively influence primary care career choice among medical students.[5] In our study, 92% of the schools had a family medicine department. Fulfillment of the deans' desire to increase collaboration and integration of services among the primary care disciplines would likely improve the impact primary care departments have on curricula, research, and leadership within their institutions.[22,23] Given the limited resources available to generalist departments and divisions, encouraging an interdisciplinary approach beyond the existing amount of coordination and integration described by most deans in our study may promote primary care and its impact in academic health centers.

In our study, deans reported that their schools had attempted to strengthen their primary care departments and divisions by placing major emphases on improving training in primary care and providing financial support for generalist departments and divisions. These efforts are concordant with recommendations made by an AAMC-sponsored advisory panel on methods to sustain the development of primary care in academic medicine.[24] They include medical school efforts to adopt curricula centering on generalist competencies and assuring adequate financial support for the generalist education of medical students.

Our study is potentially limited by self-reported data; however, self-report is appropriate in a study of deans' attitudes and perceptions of the state of primary care at their institutions and nationally. A potential response bias also exists because deans were aware that the study focused on issues relevant to academic primary care. The effect of this bias was likely limited, however, a fact exemplified by the deans' unfavorable evaluations of a number of features of academic primary care, including its research, institutional influence and prestige, and finances. The deans, on average, gave higher evaluation scores to general internal medicine compared with general pediatrics and family medicine, perhaps reflecting, in part, the high percentage of deans trained as internists and the few with general pediatrics or family medicine backgrounds. However, when adjusting for the deans' professional background, no differences were found (data not shown). Finally, because we used mostly close-ended questions, we were able only to assess the utilization of measures to strengthen academic primary care that we had identified a priori. Thus, we are unaware of unique measures used at specific institutions.

In summary, medical school deans have a generally positive attitude toward and assessment of academic primary care and believe that primary care faculty at academic health centers are central to the development of the future primary care workforce. Nationally, primary care departments and divisions remain dependent on a combination of institutional, state, and federal support to accomplish their academic mission. Preserving this support, particularly the Title VII budget, is essential. Institutions and policymakers need to study additional interventions that will maintain primary care educational efforts and improve the financial status and research activities of their primary care faculties.

References

1. Schatz IJ, Realini JP, Charney E. Family practice, internal medicine, and pediatrics as partners in the education of generalists. Acad Med. 1996;71:35–9.
2. Levinson W, Rubenstein A. Mission critical: integrating clinician-educators into academic medical centers. N Engl J Med. 1999;341:840–3.
3. Friedman RH, Alpert JJ, Green LA. Strengthening academic generalist departments and divisions. J Gen Intern Med. 1994;9:S90–S98.
4. Henderson MC, Hunt DK, Williams JW. General internists influence students to choose primary care careers; the power of role modeling. Am J Med. 1996;101:648–53.
5. Martini CJM, Veloski JJ, Barzanksy B, Xu G, Fields SK. Medical school and student characteristics that influence choosing a generalist career. JAMA. 1994;272:661–8.
6. Block SD, Clark-Chiarelli N, Singer JD. Mixed messages about primary care in the culture of U.S. medical schools. Acad Med. 1998;73:1087–94.
7. Rivo ML, Jackson DM, Clare I. Comparing physician workforce reform recommendations. JAMA. 1993;270:1083–4.
8. Association of American Medical Colleges. AAMC Policy on the generalist physician. Acad Med. 1993;68:1–6.
9. Council on Graduate Medical Education. First report of the Council. Vol. I. Washington, DC: DHHS, 1988.
10. Relman AS. Reforming the health care system. N Engl J Med. 1990;323:991–2.
11. Council on Graduate Medical Education. Fourteenth report: COGME physician workforce policies: recent developments and remaining challenges in meeting goals. Washington, DC: DHHS, Health Resources and Services Administration, 1999.
12. Peters AS, Clark-Chiarelli N, Block SD. Comparison of osteopathic and allopathic medical schools' support for primary care. J Gen Intern Med. 1999;14:730–9.

13. Haggerty RJ, Sutherland SA. The academic general pediatrician: is the species still endangered? Pediatrics. 1999;104:137–42.
14. Mainous AG, Hueston WJ, Ye X, Bazell C. A comparison of family medicine research in research intense and less intense institutions. Arch Fam Med. 2000;9:1100–4.
15. Friedman RH, Pozen JT. The academic viability of general internal medicine: the views of department of medicine chairs. Ann Intern Med. 1985;103:439–44.
16. Campbell EG, Weissman JS, Moy E, Blumenthal D. Status of clinical research in academic health centers: views fromthe research leadership. JAMA. 2001;286:800–6.
17. Politzer RM, Horab S, Fernandez E, Gamliel S, Kahn N, Mullan F. The impact of Title VII departmental and predoctoral support on the production of generalist physicians in private medical schools. Arch Fam Med. 1997;6: 531–5.
18. Fryer GE Jr, Meyers DS, Krol DM, et al. The association of Title VII funding to departments of family medicine with choice of physician specialty and practice location. Fam Med. 2002;34:436–40.
19. Pugno, PA, director, Division on Medical Education, American Academy of Family Physicians. Personal communication, December 4, 2003.
20. Block SD, Clark-Chiarelli N, Peters AS, Singer JD. Academia's chilly climate for primary care. JAMA. 1996;276:677–82.
21. Graber DR, Bellack JP, Musham C, O'Neil EH. Academic deans' views on curriculum content in medical schools. Acad Med. 1997;72:901–7.
22. Kahn N, Davis A, Wilson M, et al. The Interdisciplinary Generalist Curriculum (IGC) Project: an overview of its experience and outcomes. Acad Med. 2001;76: S9–S12.
23. Scherger JE, Rucker L, Morrison EH, Cygan RW, Hubbell FA. The primary care specialties working together: a model of success in an academic environment. Acad Med. 2000;75:693–8.
24. Inui TS, Williams WT Jr, Goode L, et al. Sustaining the development of primary care in academic medicine. Acad Med. 1998;73: 245–57.

Teaching and Learning Moment

PERSPECTIVE

As the coordinator and trainer for the University of Louisville's Standardized Patient (SP) Program, I am responsible for ensuring that all of our SP encounters produce reliable, valid outcomes, both for students on our main campus in Louisville and for those at our rural campus in Madisonville, Kentucky.

Our most recent trip to Madisonville occurred the day before Thanksgiving. The entire university closed at noon for the holiday and everyone eagerly left work to begin their holiday weekend—everyone, that is, except four standardized patients and me. We climbed into a van and headed to the Madisonville campus, two very long hours away from home, with nothing to look forward to but wet roads, insane holiday traffic, and miles and miles of damp countryside.

Once in Madisonville, I got involved with the students, although my thoughts kept drifting back to Louisville and wishing to already be home with my family. During a break in the schedule, a fourth-year medical student recognized me and stopped to talk. He told me that he would never forget a pediatric SP case that I had coordinated last year. He went on to say that during a visiting pediatric rotation in Nebraska, he had seen a toddler who presented with respiratory problems. This child was not responding to treatments, yet nothing abnormal had shown up on x-rays. Remembering the pediatric SP case, he suggested that the doctor try a flexible bronchoscope to see if the child had aspirated a foreign object. With skepticism, the physician finally agreed to the procedure, which produced a cocklebur from one of the child's lungs.

In that moment, my mood changed dramatically. Hearing his story validated my work and my missed holiday time with my family. My work, and at times sacrifices, had made an impact in this student's life. Most importantly, because of this student, a little girl in Nebraska is breathing much easier. Now, that's a new perspective.

JACKIE L. KRUGLER

Ms. Krugler is trainer and coordinator, Standardized Patient Program, Office of Curriculum Development and Evaluation, University of Louisville School of Medicine, Louisville, Kentucky.

ORIGINAL RESEARCH

The onset of HIV infection in the Leningrad region of Russia: a focus on drug and alcohol dependence

E Krupitsky,[1] E Zvartau,[1] G Karandashova,[1] NJ Horton,[2] KR Schoolwerth,[3] K Bryant[4] and JH Samet[35]

[1]St. Petersburg Scientific Research Center of Addictions and Psychopharmacology affiliated with St. Petersburg State Pavlov Medical University, St. Petersburg, Russia, [2]Department of Mathematics, Smith College, Northampton, MA, USA, Departments of [3]Medicine and [5]Social and Behavioural Sciences, Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Boston University Schools of Medicine and Public Health, Boston, MA, USA, and [4]National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health, Bethesda, MD, USA

Objectives

Within 5 years, 5 million Russians may be infected with HIV. Currently, injection drug use is the major risk factor for HIV. As Russia's alcohol consumption per capita is among the highest in the world, alcohol-associated behaviours may be an important contributor to the HIV epidemic. Our objective was to examine the prevalence of HIV infection among drug- and alcohol-dependent patients in a regional narcology hospital and in the general population in Leningrad.

Methods

All patients in the Narcology Hospital, Leningrad Regional Center of Addictions (LRCA), were tested for HIV antibody between 1997 and 2001. We reviewed these clinical records (i.e. serostatus, gender, age, and addiction) and data from the HIV/AIDS Center in the Leningrad Region (1997–2001).

Results

From 1997 to 2001, HIV prevalence at the LRCA increased from 0 to 12.7% overall, 33.4% among drug-dependent patients and 1.2% among alcohol-dependent patients. During the same 5-year period (1997–2001), 2826 persons were registered at the HIV/AIDS Center: 6, 6, 51, 780, and 1983 persons in 1997, 1998, 1999, 2000 and 2001, respectively.

Conclusions

HIV infection is exploding in the Leningrad Region, currently in injection drug users (IDUs) but potentially more broadly. The known high per capita alcohol intake in Russia heightens concern regarding the sexual transmission of HIV. Interventions to prevent such a development should include use, and assessment of the effectiveness, of known HIV prevention measures for at-risk and infected individuals.

Keywords: AIDS, alcohol and drug dependence, HIV, HIV prevalence, Russia

Received: 6 March 2003, accepted 23 July 2003

Introduction

Reports of HIV infection in Russia have increased dramatically over the past 5 years, with an estimated 300 000 HIV-infected persons in 2000 compared to 130 000 in 1999 [1]. According to forecasts, in 5 years there may be approximately 5 million HIV-infected individuals in the Russian Federation. The HIV epidemic in Russia, while perhaps initiated through sexual transmission from foreigners [2], has primarily affected the injection drug user (IDU) population [3,4]. There is concern that the HIV epidemic may expand into the general population, via sexual transmission. Experts fear that this 'second wave' of infections spread through sexual contact could transition the current drug-driven epidemic into a generalized one [5].

Injection drug use is an unequivocal major risk factor for HIV transmission. There is some evidence for an indirect role of alcohol use in HIV transmission through modulation of sexual or injection drug use-associated risk

Correspondence: Dr Jeffrey H. Samet, Section of General Internal Medicine, Clinical Addiction Research and Education (CARE) Unit, 91 East Concord St., Suite 200, Boston Medical Center, Boston, MA 02118, USA. Tel: (617) 414 7399; fax: (617) 414 4676; e-mail: jsamet@bu.edu

behaviour [6–9]. Specifically, studies have shown an association between alcohol use and unsafe sex behaviour. Stall *et al.* found, in a cohort of gay men in San Francisco, that alcohol consumption during sexual activity, at any level, was associated with high-risk sexual behaviour [10]. Rees *et al.* reported an increase in higher-risk sexual behaviour with higher levels of alcohol use among a cohort of 354 drug users [11]. Additionally, a study of 196 IDUs showed that alcohol abusers are more likely to engage in high-risk drug use [12].

The impact of alcohol use on high-risk sexual and drug use behaviours in the Russian population is important but has received limited attention. In Russia, where alcohol consumption per capita is among the highest in the world [13], HIV infection is undergoing an epidemic spread [4]. In the present study carried out in the Leningrad Region of Russia, we examined the prevalence of HIV infection among drug-dependent and alcohol-dependent inpatients in a substance abuse treatment facility and diagnosed HIV infection in the general population.

Methods

Location

The Leningrad Region of Russia surrounds St. Petersburg, located in north-west Russia, bordering Finland and Estonia. The population of this region is approximately 1.67 million, of which two-thirds dwell in urban environments and one-third in rural areas [14]. The Leningrad Regional Center of Addictions (LRCA), a 300-bed hospital, is the major referral treatment facility for alcohol- and drug-dependent individuals in this region. The beds at the LRCA are exclusively for individuals with alcohol or drug dependence as their primary diagnosis.

Data collection

All individuals who received care at the LRCA between 1997 and 2001 were tested for HIV antibody as part of the required initial assessment. The test used was the immunofluorescent antibody (IFA) test, and those samples that were positive underwent confirmation with western blot testing. We reviewed the clinical records from the LRCA, and recorded the HIV serostatus of all patients admitted during these years. Those patients hospitalized repeatedly within 1 year were only counted once; known HIV-infected patients were not retested if re-admitted in subsequent years. We also collected the following demographic information: gender, age, and primary and secondary addiction diagnoses. The diagnostic categories were drug or alcohol dependence, as assigned by the attending psychiatrist and recorded in the medical record. We specifically sought any record of drug abuse or injection drug use in the hospital record of patients identified as HIV-infected and alcohol-dependent as a primary addiction diagnosis.

We also reviewed data from the HIV/AIDS Center in the Leningrad Region for the years 1997–2001. This Center collects surveillance data on all reported cases of HIV from all hospitals including those who specialize in tuberculosis, infectious disease and narcology as well as and other testing sites in the Leningrad Region, including the LRCA. The data obtained from the HIV/AIDS Center represent the best available information about HIV prevalence in the overall Leningrad Region. The Institutional Review Boards of both the Boston Medical Center, Boston, MA and St. Petersburg Pavlov State Medical University, St Petersburg, Russia, approved this protocol.

Statistical analysis

The statistical analysis involved calculation of the prevalence of HIV diagnosis, and exact 95% confidence intervals (CIs) for this proportion. In addition, we calculated HIV prevalence stratified by primary addiction diagnosis. All analyses were carried out using SAS STAT® Software (SAS Institute, Cary, NC) [15].

Results

Leningrad Regional Center of Addictions (LRCA)

During the years 1997–2001, 8056 patients entered the LRCA. The primary diagnosis was drug dependence in 30.5% (2460) of cases and alcohol dependence in 69.5% (5595). Heroin was the drug of choice for the vast majority of drug-dependent patients. No diagnosis was available for one patient. The mean age of patients was 36.0 years; 22.6 years [standard deviation (SD) = 5.6 years] among drug-dependent patients and 41.8 years (SD = 10.6 years) among alcohol-dependent patients. Overall, 10.7% (862/8056) of patients were female. Women comprised 12.9% of drug-dependent diagnoses and 9.6% of alcohol-dependent diagnoses. All patients were tested for HIV antibody. In total, 4.8% of patients (387/8056) tested positive for HIV antibody, of which 11.1% were female. HIV prevalence stratified by diagnoses is examined by year of hospitalization and displayed in Table 1.

In 1997 and 1998, no patient, either alcohol- or drug-dependent, was HIV-infected. In 1999, 0.4% of patients (95% CI = 0.18–0.91%; seven of 1585) tested HIV antibody positive and all had a primary diagnosis of drug dependence. In 2000, 7.1% of patients (95% CI = 6.0–8.3%;

Table 1 HIV prevalence among hospitalized patients in the Leningrad Regional Center of Addictions in 1997–2001 ($n = 8055$)

	Number HIV-infected/total number of patients	
Year	Drug-dependent	Alcohol-dependent
1997	0/195	0/1014
1998	0/402	0/1010
1999	7/491 (1.4%)	0/1094
2000	126/686 (18.4%)	10/1242 (0.8%)
2001	229/686 (33.4%)	15/1235 (1.2%)

Table 2 Reported HIV cases in the Leningrad Region of Russia (1997–2001) stratified by registered risk behaviour as per the HIV/AIDS Center of the Leningrad Region

	Number HIV-infected		
Year	Drug-dependent	Alcohol-dependent	Other
1997	4	0	2
1998	4	0	2
1999	48	0	3
2000	707	9	64
2001	1796	21	166

136 of 1928) were HIV-infected. In 2001, 12.7% (95% CI = 11.2–14.2%; 244 of 1921) received a positive HIV antibody test. Between the years 1999 and 2001, the prevalence of HIV infection in drug-dependent patients increased dramatically. From 1.4% in 1999, the prevalence jumped by 17% in 2000 and by an additional 15% in 2001. Among alcohol-dependent patients, the first evidence of HIV infection was noted in 2000 and increased to 1.2% in 2001. Assessment of the hospital records of these HIV-infected alcohol-dependent patients revealed no evidence of past injection drug use. In addition, we found no secondary diagnoses of drug dependence for these patients with both HIV infection and alcohol dependence.

HIV/AIDS Center data

During the same 5-year period (1997–2001), 2826 new diagnoses of HIV infection were registered at the HIV/AIDS Center in the Leningrad Region: six in 1997, six in 1998, 51 in 1999, 780 in 2000, and 1983 in 2001. In 2000, 23.1% were female; the mean age for IDUs ($n = 707$) was 22.2 years (SD = 5.1 years), and that for alcoholics ($n = 9$) was 45.1 years (SD = 12.3 years). In 2001, 28.6% were female; the mean age for IDUs ($n = 1796$) was 22.6 years (SD = 5.0 years), and that for alcoholics ($n = 21$) was 41.7 years (SD = 9.0 years). Table 2 further describes HIV prevalence in the Leningrad Region stratified by primary diagnosis as reported to the AIDS Center.

Discussion

According to a recent UNAIDS report, HIV incidence in Eastern Europe is increasing faster than in any other part of the world [4]. This study is unique in that data on HIV prevalence are reported in a high-risk population prior to the arrival of the epidemic, at its first detection, and upon its earliest epidemic spread. The recognition of HIV infection at such an early stage in an epidemic provides an unusual opportunity to limit its spread.

Awareness of an emerging crisis with clear risks for expansion and of proven interventions for prevention and treatment potentially enables the problem to be effectively addressed. HIV infection in Russia, currently spread by injection drug use, requires all effective efforts to be employed to minimize further transmission. Underused measures include currently proscribed opioid replacement therapies (e.g. methadone), extensive outreach and needle exchange programmes, and broad public education. Whether viral transmission from the IDU population to the heterosexual population has occurred extensively is not discernible from the data presented. The early suggestion of an emergence of HIV infection in the alcohol-dependent population is worthy of close attention. Whether injection drug use, sexual transmission, or other transmission risk factors account for the alcohol-dependent subjects in this report with HIV infection is not known. We cannot definitively exclude the possibility that some misclassification may have occurred among the HIV-infected alcohol-dependent LRCA patients. However, injection drug use seems an unlikely possibility as no evidence of injection drug use or any secondary diagnosis of drug dependence was found in any of these patients. Additionally, the demographic characteristics of the alcohol- and drug-dependent groups are quite distinctive, with the alcoholic patients being much older. Concern about a higher likelihood of sexual transmission of HIV among alcohol-dependent persons has been previously raised as a result of documented increased sexual risk behaviours in this population [16,17].

One counterintuitive implication of the observation of the recent onset of HIV infection in Russia is that the Russia of 2003 has yet to experience the full impact of the ravages of this untreated disease. As the natural history from initial HIV infection to clinical AIDS is a median of 11 years, Russians are experiencing increasing HIV prevalence but not, as yet, extensive AIDS morbidity and mortality. This is unlike the past initial observations of HIV infection in many other parts of the world. This phenomenon could mislead the policy and political leaders of a country

experiencing such a scenario into a serious misconception that the issue does not require urgent attention in the midst of many other serious public health, political, and economic issues. The results from the Leningrad Region provide a stark message: HIV infection is exploding on to the scene, currently in IDUs but potentially much more broadly. In the Leningrad Region, HIV testing has been carried out aggressively in medical settings; however, it is likely that those HIV-infected individuals without such medical exposure may not become aware of their HIV diagnosis for years, resulting in substantial underestimates of the regional prevalence of HIV [18]. This perspective underscores the critical nature of efforts to address HIV prevention. Interventions for the prevention of broad transmission of HIV should include extensive distribution of condoms, sex education in public domains, implementation of harm reduction measures for injection drug users, and a serious assault on binge alcohol use.

Russia now faces numerous serious public health perils, including HIV infection and alcohol dependence. At present, the number of HIV-infected, alcohol-dependent persons is insufficient to determine whether alcohol use will contribute to the HIV epidemic in Russia. The stark increase of HIV infection in drug-dependent persons and its first occurrence in alcohol-dependent patients argues for immediate action with injection drug users and close observation of the impact of alcohol on the Russian epidemic. HIV prevention efforts in Russia will be critical in the years ahead, and interventions among substance-dependent persons in treatment could be one of many important intervention strategies.

Acknowledgements

We would like to thank the National Institute on Alcohol Abuse and Alcoholism for supporting this project, as well as GN Burakova, TY Slavina and NM Titova for their technical assistance with access to records at the Leningrad Regional Center of Addictions and the HIV/AIDS Center of the Leningrad Region. We would also like to acknowledge Jacqueline Savetsky for her editorial work on this manuscript.

Support for this study came from the National Institute on Alcohol Abuse and Alcoholism (NIAAA) of the NIH: Supplement to RO1-AA11785 (Alcohol Use, HIV Infection, and HIV Risk Behaviours in St. Petersburg, Russia).

References

1 Stephenson J. HIV/AIDS surging in Eastern Europe. *J Am Med Assoc* 2000; **284**: 3113–3114.

2 Kozlov AP, Volkova GV, Malykh AG *et al*. Epidemiology of HIV infection in St. Petersburg, Russia. *J AIDS* 1993; **6**: 208–212.

3 Dehne KL, Pokrovskiy V, Kobyshcha Y *et al*. Update on the epidemics of HIV and other sexually transmitted infections in the newly independent state of the former Soviet Union. *AIDS* 2000; **14** (Suppl. 3): S75–S84.

4 UNAIDS. *AIDS Epidemic Update*. Geneva, UNAIDS, 2001.

5 UNAIDS. *UNAIDS Calls for Intensified Support as Russia's AIDS Epidemic Spreads. Press Release*. Geneva, UNAIDS, 16 November 2000. http://www.unaids.org/whatsnew/press/eng/pressarc00/moscow161100.html

6 Samet JH. Alcohol and HIV infection. *Alcoscope Int Rev Alcoholism Management* 2000; **3**: 2–4.

7 NIAAA NIH. *Alcohol and AIDS, A Guide to Research Issues and Opportunities*. Bethesda, NIH, June 2001.

8 Kresina TF, Flexner CW, Sinclair J *et al*. Alcohol use and HIV pharmocotherapy. *AIDS Res Hum Retroviruses* 2002; **18**: 757–770.

9 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.

10 Stall R, McKusick L, Wiley J *et al*. Alcohol and drug use during sexual activity and compliance with safe sex guidelines for AIDS: The AIDS Behavioral Research Project. *Health Educ Quart* 1986; **13**: 359–371.

11 Rees V, Saitz R, Horton NJ, Samet JH. Association of alcohol consumption with HIV sex- and drug-risk behaviors among drug users. *J Sub Abuse Treatment* 2001; **21**: 129–134.

12 Stein MD, Hanna L, Natarajan R *et al*. Alcohol use patterns predict high-risk HIV behaviors among active injection drug users. *J Sub Abuse Treatment* 2000; **18**: 359–363.

13 Nemtsov AV. Estimates of total alcohol consumption in Russia, 1980–94. *Drug Alcohol Depend* 2000; **58**: 133–142.

14 Leningrad Region Online Guide. Leningrad region geography. www.cspi.org.ru/test/geo.htm (3/26/2002).

15 SAS Institute. *SAS STAT® Software. Changes and Enhancements, Release 8.2*. Cary, NC, SAS Institute Inc., 2001.

16 Avins AL, Woods WJ, Lindan CP *et al*. HIV infection and risk behaviors among heterosexuals in alcohol treatment programs. *J Am Med Assoc* 1994; **271**: 515–518.

17 Mahler J, Yi D, Sacks M *et al*. Undetected HIV infection among patients admitted to an alcohol rehabilitation unit. *Am J Psychiatry* 1994; **15**: 439–440.

18 Samet JH, Freedberg KA, Savetsky JB, Sullivan LM, Stein MD. Understanding delay to medical care for HIV infection: the long-term non-presenter. *AIDS* 2001; **15**: 77–85.

Assessing Missed Opportunities for HIV Testing in Medical Settings

Rebecca V. Liddicoat, MD, MPH, Nicholas J. Horton, ScD, Renata Urban, BA, Elizabeth Maier, BA, Demian Christiansen, MPH, Jeffrey H. Samet MD, MA, MPH

BACKGROUND: **Many HIV-infected persons learn about their diagnosis years after initial infection. The extent to which missed opportunities for HIV testing occur in medical evaluations prior to one's HIV diagnosis is not known.**

DESIGN: **We performed a 10-year retrospective chart review of patients seen at an HIV intake clinic between January 1994 and June 2001 who 1) tested positive for HIV during the 12 months prior to their presentation at the intake clinic and 2) had at least one encounter recorded in the medical record prior to their HIV-positive status. Data collection included demographics, clinical presentation, and whether HIV testing was recommended to the patient or addressed in any way in the clinical note. Prespecified triggers for physicians to recommend HIV testing, such as specific patient characteristics, symptoms, and physical findings, were recorded for each visit. Multivariable logistic regression was used to identify factors associated with missed opportunities for discussion of HIV testing. Generalized estimating equations were used to account for multiple visits per subject.**

RESULTS: **Among the 221 patients meeting eligibility criteria, all had triggers for HIV testing found in an encounter note. Triggers were found in 50% (1,702/3,424) of these 221 patients' medical visits. The median number of visits per patient prior to HIV diagnosis to this single institution was 5; 40% of these visits were to either the emergency department or urgent care clinic. HIV was addressed in 27% of visits in which triggers were identified. The multivariable regression model indicated that patients were more likely to have testing addressed in urgent care clinic (39%), sexually transmitted disease clinic (78%), primary care clinics (32%), and during hospitalization (47%), compared to the emergency department (11%), obstetrics/gynecology (9%), and other specialty clinics (10%) ($P < .0001$). More recent clinical visits (1997–2001) were more likely to have HIV addressed than earlier visits ($P < .0001$). Women were offered testing less often than men ($P = .07$).**

CONCLUSIONS: **Missed opportunities for addressing HIV testing remain unacceptably high when patients seek medical care in the period before their HIV diagnosis. Despite improvement in recent years, variation by site of care remained important. In particular, the emergency department merits consideration for increased resource commitment to facilitate HIV testing. In order to detect HIV infection prior to advanced immunosuppression, clinicians must become more aware of clinical triggers that suggest a patient's increased risk for this infection and lower the threshold at which HIV testing is recommended.**

KEY WORDS: **multiple informants; delay; HIV screening; AIDS; risk factors.**

J GEN INTERN MED 2004;19:349–356.

Received from the Clinical Addiction Research and Education (CARE) Unit (RU, EM, JHS), Section of General Internal Medicine and the Clinical A1DS Program, Departments of Medicine (RU, EM, JHS) and Social and Behavioral Sciences (JHS), and Data Coordinating Center (DC), Boston University Schools of Medicine and Public Health, Boston, Mass; and Department of Mathematics (NJH), Smith College, Northampton, Mass; and General Medicine Division (RVL), Department of Medicine, Massachusetts General Hospital, Boston, Mass.

Address correspondence and requests for reprints to Dr. Samet: Section of General Internal Medicine, 91 East Concord St., Suite 200, Boston Medical Center, Boston, MA 02118 (e-mail: jsamet@bu.edu).

Approximately 900,000 persons in the United States are infected with HIV, a national prevalence of 0.3%.[1,2] Current Centers for Disease Control and Prevention estimates suggest that 30% of these individuals are unaware of their diagnosis; as many as 275,000 people are infected with HIV but do not know it.[3] HIV infection has become a chronic and treatable disease, but in order to benefit from treatments, infected individuals must be tested for the virus and linked to medical care. This is a challenging goal, as the infection is often asymptomatic until the onset of opportunistic infections, and without treatment the median time between diagnosis and the development of clinical AIDS is 9 to 11 years.[4,5] Patients frequently become aware of their HIV serostatus very late in the course of their disease, many years after seroconversion.[6] In two different clinical studies at our institution assessing patients between 1990 and 1991, and again from 1994 to 1996, the median CD4 cell count at the time of presentation for medical care for HIV infection was 300 and 280 cells/μl, respectively.[7,8] Although delays do occur between HIV testing and linking to care,[9,10] the greatest delay occurs between initial infection and HIV testing.[7] The value of HIV testing goes beyond enabling medical care for the infected individual. Some studies have demonstrated that knowledge of HIV serostatus, particularly when positive, decreases behavior that can result in HIV transmission.[11–13] Over a decade has passed since an early call to action about HIV testing was prominently stated, "the nation's physicians and other health care providers should assume a much more active role in promoting HIV testing."[14]

In order to assess the clinical response to this call to action, we investigated the extent to which physicians in a variety of care settings within a single urban medical center adopted recommended approaches to HIV testing. We assessed 2 issues among recently tested HIV-infected patients: 1) whether prior to HIV diagnosis, clinical opportunities for health providers to recommend HIV testing had arisen; and 2) the clinician's response regarding HIV testing when a clinical opportunity occurred.

Table 1. Triggers for HIV Testing Categorized by the Likelihood of Its Clinical Association with an HIV Diagnosis

Category 1 Unequivocable Triggers	Category 2 Strongly Suggestible Triggers	Category 3 Reasonable Triggers	Category 4 Borderline Triggers
Men Sex with Men	Tuberculosis	STDs	Alcohol abuse
IDU	Varicella Zoster	Gonorrhea	Alcohol withdrawal
PCP	Lymphadenopathy	PID	Homelessness
Esophageal Candidiasis	Hepatitis B/C	Chlamydia	Psychiatric diagnosis
MAC		Syphilis	Pregnancy
Toxoplasmosis		Trichomoniasis	Abnormal Pap smear
Cryptococcemia		Genital herpes	Candida Vaginalis
Kaposi's Sarcoma		Condyloma Acuminata	Comm acquired pneumonia
Oral Thrush		Pediculosis Pubis	Otitis Media
ITP		Urethritis	Oral herpes
Bacterial endocarditis		Prostatitis	Onychomycosis
Leukopenia		Epididymitis	Unspecified HSV
Thrombocytopenia		Heroin/crack/ substance abuse	Abscess
Pancytopenia		Sepsis	Cellulitis
Parotid tumor		Perleche	Psoriasis
		Candida groin rash	Seborrheic dermatitis
		Anal/penile Candida	Sinusitis
		Meningitis (bacterial/viral)	Pyelonephritis
		Staph Aureus bacteremia	Aspergilloma

IDU, injection drug use; PCP, Pneumocystis carinii pneumonia; MAC, Mycobacterium avium complex; ITP, idiopathic thrombocytopenic purpura; STDs, sexually transmitted diseases; PID, pelvic inflammatory disease.

METHODS

Study Subjects

We performed a retrospective chart review of patients who initiated HIV-related medical care at Boston Medical Center in the HIV Diagnostic Evaluation Unit (DEU) between January 1994 and June 2001. The DEU is a weekly clinic designed for the initial assessment and triage of all nonpregnant patients presenting for primary care (PC) for their HIV infection, regardless of insurance status.[15] Referrals to the DEU come from a wide variety of sources, including inpatient hospital services, hospital outpatient clinics, the emergency department (ED), the urgent care clinic (UCC), community health centers, drug treatment programs, HIV testing sites, local correctional institutions, as well as self-referrals.

Patients were eligible for this study if they were 18 years or older, had their initial positive HIV test within 1 year of their DEU visit, and had received medical care at Boston Medical Center prior to their first positive HIV test. We included clinic visits only after March 1985, when HIV testing became widely available.[6] Medical care at Boston Medical Center was required as these were the medical records available for review. This study was approved by the Institutional Review Board of Boston Medical Center.

Data Collection

We requested medical records of patients who met inclusion criteria based on the DEU clinic's log, which contained date of initial HIV-positive test and the hospital's computerized records, which listed the dates and sites of prior clinic visits and hospitalizations. Information abstracted from medical records included the following: date of birth, gender, race/ethnicity, homelessness status prior to initial positive HIV test, date of positive HIV test, initial DEU visit date, total number of visits to Boston Medical Center in the 10 years prior to the patient's initial positive HIV test, and CD4 cell counts.[16] If 2 initial CD4 counts were available from the first month of HIV-related care, then the mean number was recorded. If, in the encounter note, a patient was noted as homeless or living in a shelter, then the patient was considered homeless. When medical records were incomplete (e.g., missing certain volumes of a multivolume chart), whatever visit data were available were assessed. Each visit in the 10 years prior to the patient's initial positive HIV test, but none earlier than March 1985, was reviewed for patient characteristics or conditions considered as clinical triggers for HIV testing (Table 1). Three of the authors—two medical students (EM, RU) and a medical resident (RL)—performed all chart reviews and

recorded if the patient had any of the listed characteristics or conditions. Charts were reviewed with a predetermined list of clinical conditions. Chart reviewers could list other conditions that they considered potential triggers for HIV testing; however, after review with study investigators, greater than 95% of triggers came from the predetermined list. The site of the clinical encounter was recorded.

Triggers for HIV Testing

Triggers were hierarchically categorized based on the level of clinical suspicion for HIV infection associated with each after review of the medical literature.[16] For example, injection drug use (IDU) and men having sex with men (MSM) were defined as category 1 (unequivocable triggers), while homelessness and alcohol abuse were defined as category 4 (borderline triggers). Visits were categorized according to the highest (i.e., most unequivocable) category trigger present during that visit. For example, if a visit noted injection drug use, homelessness, lymphadenopathy, and gonorrhea, it was defined as a category 1 visit, due to the presence of IDU. We stratified triggers as either unequivocal, strongly suggestible, reasonable, or borderline for triggering HIV testing discussions; categories were identified as 1 to 4, respectively. If a trigger diagnosis was considered but not made definitively such as "tuberculosis versus bronchitis," then that was labeled borderline, category 4, even if the differential diagnosis included conditions listed in categories 1 to 3.

A "missed opportunity" for HIV testing was a visit in which HIV was not discussed yet contained at least 1 of the patient characteristics defined as "triggers." Visits with physicians, nurse practitioners, nurses, and dentists were reviewed for the presence of HIV trigger conditions recorded in the clinical note. Although clinician type was not possible to abstract in the chart review, physicians account for the overwhelming majority of clinical encounters with the exception that nurse and nurse practitioner encounters were most common in the sexually transmitted disease (STD) clinic. If a visit included 1 or more triggers, then additional information was collected about that visit assessing visit location, date, and other triggers, as well as whether HIV testing was considered or recommended to the patient. Examples of the "HIV considered" group were notes that stated "HIV negative 4 months ago" and "HIV a possibility." All potential HIV-associated conditions were recorded and categorized into the diagnoses in Table 1.

Defining Patient Characteristics

Prior to medical record review, we explicitly defined several patient characteristics. Alcohol abuse was recorded if "alcohol abuse" or "alcohol withdrawal" was noted or if the patient had a history of alcohol detoxification or was being admitted to alcohol detoxification. If a patient was recorded as drinking 12 or more beers or 1 pint of liquor/day, that was also recorded as alcohol abuse. Psychiatric diagnoses included depression and anxiety in addition to schizophrenia and psychosis. Abnormal Pap smears included histories and/or diagnoses of cervical dysplasia and cervical cancer. Substance abuse was checked if it was not specified which substances were being abused. Tuberculosis was not recorded if a patient had only a positive PPD in the absence of a positive chest radiograph. Lymphadenopathy included lymph node biopsy in addition to generalized and localized lymphadenopathy found on physical exam. Pregnancy included those patients admitted for spontaneous or induced abortion in addition to those that were receiving prenatal care or were admitted for delivery. Homelessness was recorded for persons living in shelters in addition to street dwellers. The time period of the clinical encounter was divided into 4-year intervals beginning with 1985 to 1988 and continuing through 1997 to 2001.

Statistical Analysis

Simple proportions of patients who had a record of a clinician's recommending or considering HIV infection were calculated for each of the trigger conditions. We examined triggers associated with a provider's missed opportunity to consider HIV testing using generalized estimating equation (GEE)[17,18] as implemented in SAS PROC GENMOD (SAS Institute, Cary, NC). These models accounted for the correlation between visits recorded for a given subject. An exchangeable working correlation matrix was assumed, and an empiric variance estimator was used to generate standard error estimates for the regression parameters. Potential confounding variables (gender, age of first diagnosis, race/ethnicity, homeless status, and time period of the clinical encounter) were included in the model, along with visit site and trigger category. In secondary analyses, the model was also fit excluding visits from 1985 to 1993, to assess the consistency of results only using more recent data, and also refit excluding category 4 (borderline triggers).

We performed additional secondary analyses with 10 selected triggers of particular clinical interest: men having sex with men, community-acquired pneumonia, hepatitis B and/or C, homelessness, sexually transmitted diseases, weight loss, zoster, injection drug use, cocaine use, and alcohol abuse. For the analyses of the 10 individual triggers, we used an extended GEE model considering whether use of a particular trigger will yield different associations. These methods are similar to fitting separate logistic regression models for each of these specified triggers. A limitation of fitting separate models has been the lack of a method for deciding whether regression coefficients are different in the separate models, and how to combine results if they are not. We used new techniques[19–21] to fit models that allow different parameters for the association between trigger and outcome to appear in each of the equations. Using this model, it is possible to test whether there are different associations between each of these triggers and the outcome. Each subject contributes 10 observations to the data set (one for each trigger) for each visit, with an indicator as to whether that trigger was observed for that subject. An exchangeable working correlation structure was used, with subject as the

clustering variable. For these models, using the previous results for guidance, a parsimonious regression was fit that dropped race/ethnicity, homelessness status, and used a 2 category period variable (1985 to 93 vs 1994 to 2001).

RESULTS

Between 1994 and 2001, 1,400 patients were seen at the DEU clinic and of those, 358 preliminarily met the study eligibility criteria based on DEU and administrative hospital computerized records: HIV tested in the previous 12 months and received prior care at Boston Medical Center. Of the eligible 358 patients, medical records of 256 (72%) were located by the hospital records department. After review of the individual medical records, 7% (19/256) of patients were excluded because they did not meet 1 of the 2 eligibility criteria. Among the 237 remaining patient charts, we reviewed a total of 3,742 clinic visits. An additional 16 medical records were excluded because they did not have visits after March 1985, the year that the HIV test became widely available.[6] Among the remaining 221 patients' medical records, 5 were incomplete, but the available data were included in this study. There were 3,424 clinic visits for the 221 patients included in the final analyses.

All 221 patients had one or more triggers for HIV testing found in at least one encounter note. Triggers for HIV testing were noted in 50% (1,702/3,424) of the eligible visits reviewed among the 221 patients. HIV testing was recommended to the patient in 18% (299/1,702) of visits in which triggers were noted. HIV was considered in the note by the clinician without recommending testing in another 10% of visits (169/1,702). In total, HIV testing was recommended or considered in the provider note in 27% (468/1,702) of visits with triggers noted. The median number of visits per patient with a trigger was 5 (mean 7.7). The median number of triggers that a patient had per visit was 2.0 (mean 2.1).

Table 2. Characteristics of HIV-infected Patients Who Received Medical Care at Boston Medical Center Prior to Their HIV Diagnosis (*N* = 221)

	Characteristics	*n* (%)
Race/ethnicity	African-American	109 (49)
	White	27 (12)
	Hispanic	28 (13)
	Haitian/African	50 (23)
	Other	7 (3)
Age,* y†	18 to 24	10 (5)
	25 to 34	68 (31)
	35 to 44	96 (43)
	45 to 54	36 (16)
	55+	10 (5)
Gender	Male	146 (66)
	Female	75 (34)
CD4* (cells/μl)	≥200	124 (56)
	<200	96 (44)

* N = *220*.
† *Age at time of DEU clinic presentation.*

Demographic characteristics of the 221 patients (Table 2) include the following: 66% male, 49% African-American, 23% immigrants from an HIV endemic country, and 22% homeless. The mean age at the time of a positive HIV test was 39 years. In 44% of patients (96/220), the initial CD4 count was less than 200 cells/μl when diagnosed with HIV. The mean CD4 count was 328 cells/μl, while the median was 256 cells/μl. Only 51% (113/221) of patients had any PC visit in the Boston Medical Center system prior to the date of their initial positive HIV test.

Thirty-nine percent (670) of the clinical visits (*n* = 1,702) with HIV triggers were to the ED (370) or UCC (300). Primary care was the second most common clinical site with 18% (306). Hospitalization accounted for 13% (218) of such visits and obstetrics/gynecology 7% (119). Although HIV was addressed in 28% of the 1,702 visits, the percentage of these visits varied widely by site (Table 3). While 32% of visits to PC clinic and 39% of visits to UCC addressed HIV, only 12% of ED visits considered HIV infection. The site that most routinely considered HIV was STD clinic (78%), followed by hospitalization (47%). Other sites with low percentages for addressing the issue of HIV testing were other specialists (10%), obstetrics/gynecology (9%), and dermatology (14%).

The multivariable model for missed opportunities for recommending testing or considering HIV found that gender was a borderline significant predictor, with women being more likely to have a missed opportunity (odds ratio [OR], 1.42; 95% confidence interval [CI], 0.98 to 2.07). There was no overall association between race/ethnicity and discussion (degrees of freedom [d.f.], 4; $P = .44$). Older age at first HIV diagnosis was associated with missed opportunities (OR, 1.26 for each additional decade of age; 95% CI, 1.02 to 1.55), while homelessness ($P = .90$) had no significant association with HIV discussion or testing.

The year of the visit had a significant association with addressing HIV, showing that more HIV testing occurred over time (d.f., 3; $P < .001$). Compared to visits during the periods 1997 to 2001, visits during 1985 to 1988 (OR, 12.0; 95% CI, 6.0 to 23.9), 1989 to 1992 (OR, 3.6; 95% CI, 2.4 to 5.3), and 1993 to 1996 (OR, 1.9; 95% CI, 1.4 to 2.6) had greater odds of missed opportunities. Site of visit was also a significant predictor of missed opportunities for discussion (d.f., 7; $P < .0001$). Compared to the UCC, visits to the ED (OR, 4.2; 95% CI, 2.6 to 6.7), obstetrics/gynecology clinic (OR, 2.0; 95% CI, 1.1 to 3.6), other specialty clinic (OR, 4.0; 95% CI, 2.3 to 6.9), and surgical clinic (OR, 10.3; 95% CI, 2.0 to 53.3) had greater odds of a missed opportunity. Visits to the PC clinic (OR, 1.0; 95% CI, 0.7 to 1.5) were not significantly different from the UCC, while the STD clinic had lower odds of a missed opportunity (OR, 0.07; 95% CI, 0.04 to 0.15).

Trigger category was significantly associated with missed opportunities for testing (d.f., 3; $P < .0001$). Table 4 shows the percentage of time that HIV was discussed in visits stratified by trigger category. Compared to category 4 (borderline triggers), category 1 had lower odds of missed

Table 3. By Visit Site, the Percentage of Visits Where HIV Testing Was Recommended or Considered by a Clinician Stratified by Trigger Category

Visit Site	HIV Testing Was Recommended or Considered									
	Category 1 Trigger % (*n**/total†)		Category 2 Trigger % (*n*/total)		Category 3 Trigger % (*n*/total)		Category 4 Trigger % (*n*/total)		Total % (*n*/total)	
Primary care	67	(45/67)	42	(27/65)	22	(20/91)	7	(6/83)	32	(98/306)
ED	23	(19/84)	16	(10/64)	11	(11/104)	3	(3/118)	12	(43/370)
Urgent care center	62	(56/90)	41	(36/87)	31	(22/72)	6	(3/51)	39	(117/300)
STD clinic	100	(8/8)	89	(8/9)	74	(51/69)	100	(2/2)	78	(69/88)
Obstetrics/ gynecology	0	(0/4)	20	(1/5)	10	(5/48)	8	(5/62)	9	(11/119)
Other/ specialist	29	(8/28)	12	(3/26)	3	(1/32)	0	(0/28)	11	(12/114)
Hospital	68	(60/88)	73	(22/30)	32	(17/53)	9	(4/47)	47	(103/218)

* *n = number of HIV recommended/discussed visits.*

† *Total number of visits at that clinical site within the column's particular trigger category.*

ED, emergency department; STD, sexually transmitted disease.

opportunities (OR, 0.05; 95% CI, 0.03 to 0.08), as did category 2 (OR, 0.13; 95% CI, 0.08 to 0.21) and category 3 (OR, 0.27; 95% CI, 0.18 to 0.42). There were also statistically significant differences in missed opportunities among trigger categories 1 to 3. Compared to categories 2 and 3, respectively, category 1 had lower odds of missed opportunities (OR, 0.38, 95% CI, 0.25 to 0.58; OR, 0.18, 95% CI, 0.12 to 0.28). Category 2 had lower odds of missed opportunities than category 3 (OR, 0.49; 95% CI, 0.34 to 0.70).

In secondary analyses using only visit data from 1993 to 2001, the results were generally consistent, though women had significantly more missed opportunities for HIV testing ($P = .01$), and the age association was attenuated ($P = .16$). Results from the regression model were also similar when the category 4 trigger visits were excluded from the analysis.

When individual triggers were compared to one another, there was a significant difference between trigger type and whether HIV testing was recommended or considered (d.f., 9; $P = .05$), indicating that there were significant differences in clinicians' perceptions of the associations between the individual triggers and HIV discussion (while controlling for gender, age, location, and period). To help illustrate these differences, Table 5 lists the unadjusted percentage of visits with specific triggers where HIV was recommended or considered. Men having sex with men as a trigger was associated with the highest proportion of HIV testing being recommended or considered, 71%. When injection drug use was noted, HIV testing was recommended or considered 54% of the time. Zoster was the weakest trigger for HIV testing recommendation or consideration among the 10 individually assessed triggers.

DISCUSSION

More than 2 decades after AIDS was first described, patients continue to present for initial HIV-related medical care years after acquiring the virus. Although diagnosis of HIV infection at an asymptomatic stage is a challenge, it

Table 4. Examined by Patient Characteristic, Percentage of Visits Where HIV Testing Was Recommended, or Considered by a Clinician Stratified by Trigger Category

Patient Characteristics	HIV testing was recommended or considered									
	Cat 1 Trigger % (*n*/total)		Cat 2 Trigger % (*n*/total)		Cat 3 Trigger % (*n*/total)		Cat 4 Trigger % (*n*/total)		Total % (*n*/total)	
AA	48	(119/250)	36	(70/193)	26	(86/334)	6	(15/240)	29	(290/1017)
White	63	(45/71)	6	(4/64)	26	(12/46)	2	(1/52)	27	(62/233)
Hispanic	62	(24/39)	41	(16/39)	20	(11/56)	6	(3/51)	29	(54/185)
Haitian/African	58	(14/24)	30	(20/66)	22	(13/58)	8	(7/89)	23	(54/237)
Other	0	(0/6)	33	(3/9)	71	(10/14)	0	(0/1)	43	(13/30)
Male	49	(143/291)	32	(82/259)	29	(75/258)	4	(9/201)	31	(309/1009)
Female	60	(59/99)	38	(31/81)	23	(57/250)	6	(17/263)	24	(164/693)
All	52	(202/390)	33	(113/340)	26	(132/508)	6	(26/433)	28	(473/1702)

AA, African-American.

Table 5. Examination by Specific HIV Triggers in Medical Encounters Between 1994 and 2001 Where HIV Testing Was Recommended or Considered by the Clinician

HIV Trigger	% (Number of visits with HIV testing recommended or considered/total number of visits with triggers)
Men sex with men	71 (32/45)
Weight loss	68 (54/80)
Injection drug use	54 (91/167)
Hepatitis B and/or C	50 (51/103)
Community acquired pneumonia	50 (66/132)
Sexually transmitted disease	46 (100/217)
Crack/cocaine	42 (98/235)
Alcoholism/alcohol abuse	35 (49/139)
Homelessness	34 (57/167)
Herpes Zoster	21 (19/90)

is a goal worth pursuing, as early testing to achieve this objective can benefit both the patient and society.[13,22–25] Examination of the medical care of HIV-infected persons prior to an HIV diagnosis has received limited attention.

Demographic characteristics of the 221 patients in this study are similar to national averages for HIV-infected persons: 66% male and 49% African-American compared to 70% male and 54% African-American nationally;[26] 44% with CD4 counts below 200 cells/μl, similar to 36% found in several other urban centers.[27] Our study found that in less than 1 of 5 encounters with a clinical trigger for HIV infection, documentation was found in the chart that HIV testing was recommended to the patient. Clinical documentation of "consideration" of HIV infection, a less stringent criterion, occurred in only 28% of encounters. In the case of category 1 and 2 visits, with triggers such as injection drug use, lymphadenopathy, and varicella zoster, testing was recommended or considered only 52% for category 1 and 33% for category 2 visits.

A large number of the visits with triggers for HIV testing (39%, 670/1,702) were found in patients presenting to the ED or UCC, while 18% of visits were to PC clinic. Only half of the 221 patients had even 1 PC visit. This finding indicates that successful early HIV diagnosis in medical settings requires outreach beyond the PC clinical arena.

The site of the encounter was highly associated with HIV testing's being recommended or considered. Hospitalized and STD clinic patients had a relatively high level of addressing HIV while the ED and obstetrics/gynecology and surgical clinics had lower levels. Barriers to testing at these sites may relate to time pressures and absence of a structured system to facilitate testing. The STD clinic routinely highlights HIV testing on patient forms as a physician reminder, while the ED did not have a reminder system or a staff member to encourage HIV testing. However, neither the UCC nor PC clinic had a system for HIV testing, yet the rates of recommending HIV testing, while still poor, were 3 times that of the ED. Several possible explanations may account for the discrepancy between testing in the UCC and ED: the level of medical acuity in the UCC is less than the ED setting, and ED staff have less training in preventive medicine as compared to internal medicine physicians in the UCC. Only 21% of TB clinic patients with triggers had a documented HIV test recommendation. Considering that 28% to 46% of adults with TB in the United States are HIV infected,[28] 21% is a very low percentage. However, the small number of TB clinic encounters (n = 33) makes conclusions in this setting difficult.

Missed opportunities for HIV testing may reflect the lack of adequate HIV services in the clinics and ED. Requiring patients to return to an unfamiliar clinical setting like the ED to receive HIV test results may be more likely to result in persons not returning for test results.[29] It is also necessary for transient care sites to have effective systems for follow-up so that patients who test positive are able to engage in care.[30]

Our data indicate that even when triggers for HIV are present, clinicians either do not think of testing or do not document the consideration of this diagnosis. Developing systems whereby patients with selected "trigger" conditions presenting to EDs or other sites with high HIV prevalence are automatically offered HIV testing independent of the provider, would increase testing yet not add substantially to the burden of the provider. The Centers for Disease Control and Prevention has recommended that all high prevalence hospitals, those with greater than 1 new diagnosis of HIV per 1,000 inpatients, should implement testing for all inpatients. Implementation of such recommendations has yielded enhanced case findings of undiagnosed HIV-infected individuals.[31] Expanding this recommendation to other high-risk clinical sites would enhance testing efforts.

In November 2002, the Food and Drug Administration approved OraQuick, a rapid fingerstick test with results in 20 minutes (sensitivity of 99.6%, specificity of 100%).[32] Its utility may be particularly valuable in the UCC and ED setting in which patients are less likely to return for test results.[29] If instituted in the ED and other high-volume sites, rapid testing might enhance testing and increase linkage to care.[33,34]

The lower proportion of testing for females (24%) compared to males (31%) has been shown previously.[35] As the HIV epidemic becomes increasingly equally distributed between men and women, this past bias in provider risk perception for HIV needs to be eliminated.

Another potential barrier to expanded testing is the requirement of informed consent prior to testing. This requires a provider's time, which may decrease testing recommendation. If broader testing is to be implemented and explicit written informed consent continues to be considered essential prior to testing, then resources to enable this activity will be necessary.

A limitation of our study was dependence on chart review methodology to determine whether HIV testing had been considered or recommended between providers and patients. It is possible that discussions occurred but were not

accurately documented. Another limitation is the need to make explicit and categorize the triggers for HIV testing derived from recently published reports. This is necessarily an approximation and differences of opinion may exist about the categorization of particular characteristics or diagnoses. In addition, category 4 triggers may have been less broadly appreciated in the earlier years assessed. Another limitation is that these data reflect a single urban northeastern U.S. hospital. Although representative of many clinical settings, generalization of these results may not be applicable to all hospitals. Patients were chosen based on the existence of documented clinical encounters at the study institution prior to their knowledge of HIV infection. Some patients may have been ineligible due to their receiving care at other institutions prior to a positive HIV test. These patients' encounters may have differed from study patients.

CONCLUSIONS

Missed opportunities for earlier HIV testing have been the norm for patients who received medical care prior to their HIV diagnoses. Recommendation or consideration of HIV testing was noted in the clinical record in less than a third of such opportunities. Many HIV-infected patients received their medical care prior to HIV diagnosis at transient sites, the emergency department, and urgent care clinic. These data indicate the need to initiate expansive HIV testing in transient sites with high volumes of HIV-infected individuals. It should be the responsibility of providers in all clinical sites to ensure that opportunities to address HIV testing are not missed.

We appreciate Jennifer Whiteley's assistance with the preparation of this manuscript and Colleen LaBelle for her selfless efforts to address patients' needs in the DEU clinic and with the preparation of this manuscript. We would also like to thank Ernesto Falcon and the Medical Records Department of Boston Medical Center for their assistance with this project. Dr. Horton was supported in part by NIH grant NIMH RO1-MH54693. Ms. Urban and Ms. Maier were supported by the Boston University Medical Student Summer Research Program and the General Clinical Research Center at Boston University School of Medicine (M01-RR00533).

REFERENCES

1. McQuillan GM, Khare M, Karon JM, Schable CA, Vlahov D. Update on the seroepidemiology of human immunodeficiency virus in the United States household population: NHANES III, 1988–94. J Acquir Immune Defic Syndr Hum Retrovirol. 1997;14:355–60.
2. CDC. Guidelines for national human immunodeficiency virus case surveillance, including monitoring for human immunodeficiency virus infection and acquired immunodeficiency syndrome. MMWR Recomm Rep. 1999;48:1–31.
3. CDC. HIV and AIDS cases reported through June 2000. HIV AIDS Surveill Rep. 2000;12:1–41.
4. Bacchetti P, Moss AR. Incubation period of AIDS in San Francisco. Nature. 1989;338:251–3.
5. Morgan D, Mahe C, Mayanja B, Okongo JM, Lubega R, Whitworth JA. HIV-1 infection in rural Africa: is there a difference in median time to AIDS and survival compared with that in industrialized countries? AIDS. 2002;16:597–603.
6. Wortley PM, Chu SY, Diaz T, et al. HIV testing patterns: where, why, and when were persons with AIDS tested for HIV? AIDS. 1995;9:487–92.
7. Samet JH, Freedberg KA, Savetsky JB, Sullivan LM, Stein MD. Understanding delay to medical care for HIV infection: the long-term non-presenter. AIDS. 2001;15:77–85.
8. Samet JH, Retondo MJ, Freedberg KA, Stein MD, Heeren T, Libman H. Factors associated with initiation of primary medical care for HIV-infected persons. Am J Med. 1994;97:347–53.
9. Samet JH, Freedberg KA, Stein MD, et al. Trillion virion delay: time from testing positive for HIV to presentation for primary care. Arch Intern Med. 1998;158:734–40.
10. Turner BJ, Cunningham WE, Duan N, et al. Delayed medical care after diagnosis in a US national probability sample of persons infected with human immunodeficiency virus. Arch Intern Med. 2000;160:2614–22.
11. Berrios DC, Hearst N, Coates TJ, et al. HIV antibody testing among those at risk for infection. The National AIDS Behavior Surveys. JAMA. 1993;270:1576–80.
12. Coates TJ. Efficacy of voluntary HIV-1 counseling and testing in individuals and couples in Kenya, Tanzania, and Trinidad: a randomized trial. The Voluntary HIV-1 Counseling Testing Efficacy Study Group. Lancet. 2000;356:103–12.
13. Higgins DL, Galavotti C, O'Reilly KR, et al. Evidence for the effects of HIV antibody counseling and testing on risk behaviors. JAMA. 1991;226:2419–29.
14. Rhame FS, Maki DG. The case for wider use of testing for HIV infection. N Engl J Med. 1989;320:1248–54.
15. Samet JH, Libman H, LaBelle C, 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.
16. Freedberg KA, Samet JH. Think HIV: why physicians should lower their threshold for HIV testing. Arch Intern Med. 1999;159:1994–2000.
17. Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. Biometrics. 1986;42:121–30.
18. Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. Biometrika. 1986;73:13–22.
19. Horton NJ, Laird NM, Zahner G. Use of multiple informant data as a predictor in psychiatric epidemiology. Int J Meth Psychiatr Res. 1999;8:6–18.
20. Horton NJ, Laird NM, Murphy JM, Monson RR, Sobol AM, Leighton AH. Multiple informants: mortality associated with psychiatric disorders in the Stirling County Study. Am J Epidemiol. 2001;154:649–56.
21. Pepe MS, Whitaker RC, Seidel K. Estimating and comparing univariate associations with application to the prediction of adult obesity. Stat Med. 1999;18:163–73.
22. Stover E, Steinberg L. Early detection of HIV: implications for prevention research. J Acquir Immune Defic Syndr. 2000;25:S93.
23. Choi KH, Coates TJ. Prevention of HIV infection. AIDS. 1994;8:1371–89.
24. Ho DD. Time to hit HIV, early and hard. N Engl J Med. 1995;333:450–1.
25. Kaplan JE, Masur H, Holmes KK, et al. USPHS/IDSA guidelines for the prevention of opportunistic infections in persons infected with human immunodeficiency virus: an overview. USPHS/IDSA Prevention of Opportunistic Infections Working Group. Clin Infect Dis. 1995;21:S12–31.
26. Centers for Disease Control and Prevention. A Glance at the HIV Epidemic. Available at: http://www.cdc.gov/nchstp/od/news/At-a-Glance.pdf. Accessed July 21, 2003.
27. Dybul M, Bolan R, Condoluci D, et al. Evaluation of initial CD4+ T cell counts in individuals with newly diagnosed human immunodeficiency virus infection, by sex and race, in urban settings. J Infect Dis. 2002;185:1818–21.

28. Weis SE, Foresman B, Cook PE, Matty KJ. Universal HIV screening at a major metropolitan TB clinic: HIV prevalence and high-risk behaviors among TB patients. Am J Public Health. 1999;89:73–5.
29. Samet JH, Mulvey KP, Zaremba N, Plough A. HIV testing in substance abusers. Am J Drug Alcohol Abuse. 1999;25:269–80.
30. Batal H, Biggerstaff S, Dunn T, Mehler PS. Cervical cancer screening in the urgent care setting. J Gen Intern Med. 2000;15:389–94.
31. 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.
32. Anonymous. Advancing HIV Prevention: New Strategies for a Changing Epidemic-United States, 2003. MMWR April 18, 2003.
33. Spielberg F, Kassler WJ. Rapid testing for HIV antibody: a technology whose time has come. Ann Intern Med. 1996;125:509–11.
34. Irwin K, Olivo N, Schable CA, Weber JT, Janssen R, Ernst J. Performance characteristics of rapid HIV antibody assay in a hospital with a high prevalence of HIV infection. CDC-Bronx-Lebanon HIV Serosurvey Team. Ann Intern Med. 1996;125:471–5.
35. Stein MD, Crystal S, Cunningham WE, et al. Delays in seeking HIV care due to competing caregiver responsibilities. Am J Public Health. 2000;90:1138–40.

 doi:10.1111/j.1360-0443.2004.00670.x

Uptake and adherence to highly active antiretroviral therapy among HIV-infected people with alcohol and other substance use problems: the impact of substance abuse treatment

Anita Palepu[1], Nicholas J. Horton[5], Nicole Tibbetts[2], Seville Meli[3] & Jeffrey H. Samet[3,4]

Centre for Health Evaluation and Outcome Sciences, St Paul's Hospital, Department of Medicine, University of British Columbia, Vancouver, British Columbia, Canada,[1] DM-STAT Inc., Medford, MA, USA,[2] Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Departments of Medicine,[3] Social and Behavioral Sciences[4] and Boston University Schools of Medicine and Public Health and Department of Mathematics, Smith College, Northampton, MA, USA[5]

Correspondence to:
Dr Anita Palepu
620B-1081 Burrard Street
Vancouver, BC
Canada V6Z 1Y6
Tel: 604 682 2344/local 63194
Fax: 604 806 8005
E-mail: anita@hivnet.ubc.ca

Submitted 12 February 2003;
initial review completed 7 May 2003;
final version accepted 23 October 2003.

ABSTRACT

Aim We examined the association of substance abuse treatment with uptake, adherence and virological response to highly active antiretroviral therapy (HAART) among HIV-infected people with a history of alcohol problems.

Design Prospective cohort study.

Methods A standardized questionnaire was administered to 349 HIV-infected participants with a history of alcohol problems regarding demographics, substance use, use of substance abuse treatment and uptake of and adherence to HAART. These subjects were followed every 6 months for up to seven occasions. We defined substance abuse treatment services as any of the following in the past 6 months: 12 weeks in a half-way house or residential facility; 12 visits to a substance abuse counselor or mental health professional; or participation in any methadone maintenance program. Our outcome variables were uptake of antiretroviral therapy, 30-day self-reported adherence and HIV viral load suppression.

Findings At baseline, 59% (205/349) of subjects were receiving HAART. Engagement in substance abuse treatment was independently associated with receiving antiretroviral therapy (adjusted OR; 95% CI: 1.70; 1.03–2.83). Substance abuse treatment was not associated with 30-day adherence or HIV viral load suppression. More depressive symptoms (0.48; 0.32–0.78) and use of drugs or alcohol in the previous 30 days (0.17; 0.11–0.28) were associated with worse 30-day adherence. HIV viral load suppression was positively associated with higher doses of antiretroviral medication (1.29; 1.15–1.45) and older age (1.04; 1.00–1.07) and negatively associated with use of drugs or alcohol in the previous 30 days (0.51; 0.33–0.78).

Conclusion Substance abuse treatment was associated with receipt of HAART; however, it was not associated with adherence or HIV viral load suppression. Substance abuse treatment programs may provide an opportunity for HIV-infected people with alcohol or drug problems to openly address issues of HIV care including enhancing adherence to HAART.

KEYWORDS Access, adherence, alcohol, antiretroviral therapy, HIV, substance abuse treatment, substance use.

INTRODUCTION

Highly active antiretroviral therapy (HAART) has led to significant reductions in morbidity and mortality as well as enhanced quality of life for many HIV-infected individuals (Palella *et al.* 1998; Hogg *et al.* 1999). Certain HIV-infected subgroups, such as substance users, appear to have a lower uptake of antiretroviral therapy compared to other HIV-positive people in Canada, the United States and Europe (Celentano *et al.* 1998; Strathdee *et al.* 1998; Carrieri *et al.* 1999; Mocroft *et al.* 1999; Lucas *et al.* 2001). Among eligible HIV-infected injection drug users in Vancouver and Baltimore only 40 and 51%, respectively, reported receiving antiretroviral therapy. These studies found that not being enrolled in an addiction treatment program was associated with failure to receive medications (Celentano *et al.* 1998; Strathdee *et al.* 1998). In addition, ongoing drug use lowered the likelihood of antiretroviral prescription among HIV-infected drug users who had regular access to AIDS specialized hospital care (Carrieri *et al.* 1999; Stein *et al.* 2000; Lucas *et al.* 2001). Conversely, studies have found participation in methadone maintenance increased use of HAART and may reinforce adherence to medical recommendations (Sambamoorthi *et al.* 2000; Moreno *et al.* 2001).

Ongoing drug use appears to be associated with poorer adherence. Lucas *et al.* found that active illicit drug users had inferior self-reported adherence to HAART resulting in poorer virological and immunological outcomes, compared to former and non-drug users (Lucas *et al.* 2001). Among HIV-infected drug users in methadone maintenance, ongoing illicit drug injection was the only factor significantly associated with multiple measures of antiretroviral therapy non-adherence (Stein *et al.* 2000). Moreover, HIV-infected opioid users in buprenorphine maintenance treatment achieved higher levels of adherence than drug users not in treatment (Moatti *et al.* 2000). These data suggest that substance abuse treatment services may be key to optimizing antiretroviral adherence among drug users.

Despite the findings that alcohol problems are commonly encountered among HIV-infected people (Lefevre *et al.* 1995; Petry 1999; Samet *et al.* 2003), few data are available regarding the role of substance abuse treatment services for HIV-infected people with alcohol problems. Alcohol is a risk factor for poorer medication adherence; in the HIV Cost and Utilization of Services Study, both people with heavy alcohol and drug use were less likely to achieve 100% 7-day adherence to antiretroviral therapy (Galvan *et al.* 2002). In this study it was hypothesized that engagement in substance abuse treatment services would improve the uptake of antiretroviral therapy, adherence and HIV viral load suppression in a cohort of HIV-infected people with alcohol problems. We examined these issues among participants in the HIV-Alcohol Longitudinal Cohort (HIV-ALC) study, which includes HIV-infected people with a history of alcohol problems.

METHODS

Study design

We analyzed data from a prospective cohort of HIV-infected patients with a history of alcohol problems. One hundred and fifty-one subjects in the cohort participated in a randomized controlled trial of a HAART adherence intervention (ADHERE study); appropriate adjustments were made to the analysis to account for the trial (Samet *et al.* 2002).

Study population

Patients who were HIV-infected and had a history of alcohol problems were identified by explicit eligibility criteria. All potential subjects who gave two or more positive responses to the CAGE (Cut-down, Annoyed, Guilt, Eye-opener) questionnaire (Ewing 1984; Buchsbaum *et al.* 1991; Fiellin *et al.* 2000; Samet *et al.* 2003), a screening test for life-time alcohol problems (sensitivity ~80%, specificity ~90%), were eligible. In addition, those patients recruited from the Boston Medical Center HIV Diagnostic Evaluation Unit (DEU) (Samet *et al.* 1995) who did not meet CAGE criteria, were eligible if one of two attending physicians made a specific diagnosis of alcohol abuse or dependence. Thus subjects with alcohol problems, despite not being detected by the CAGE, were detected by the clinical interview and recruited. However, within the DEU site most subjects were recruited based on CAGE criteria. Other entry criteria included the following: fluency in English or Spanish; Mini-Mental State Examination (MMSE) score greater or equal to 21 (Folstein *et al.* 1975); and no plans to move from the Boston area in the next 2 years. As chronic alcohol use is associated with cognitive impairment, we used the MMSE cut-off of 21 to exclude subjects in whom such impairment may preclude obtaining informed consent, an accurate and complete interviewer-administered questionnaire or a follow-up interview. The Institutional Review Boards of Boston Medical Center and Beth Israel Deaconess Medical Center approved this study.

From July 1997 to July 2001, recruitment of subjects occurred by multiple methods and from several sites: Boston Medical Center HIV Diagnostic Evaluation Unit (56%); posted flyers (17%); Boston Medical Center Primary Care Clinic (13%); respite facility for homeless people (5%); methadone clinic (4%); subject referrals (4%); and Beth Israel Deaconess Medical Center (2%). The majority of subjects was recruited from medical set-

tings that addressed HIV-related issues. The eligibility criteria of a history of alcohol problems was determined by the CAGE questionnaire in 313/349 (90%) of subjects, and based on clinical assessment in 36/349 (10%) of subjects. Diagnostic interviews for alcohol problems in a sample of these subjects ($n = 141$) revealed an 80% life-time history of alcohol dependence (113/141) or abuse 15% (21/141) in over 90% (Samet *et al.* 2004).

Data collection

After obtaining informed consent, a research associate or study investigator interviewed subjects using a standardized instrument to ascertain baseline information including the following: demographics, HIV risk behaviors, alcohol severity, use of substance abuse treatment services and health care utilization in the preceding 6 months. For the Spanish interview standardized scales in Spanish were used when available; the remainder of the questionnaire was translated from English into Spanish, back-translated to check for accuracy, and then corrected. We attempted to obtain CD4 cell counts and HIV RNA (viral load) levels on all subjects. Laboratory tests performed within 6 months of the interview as part of clinical care were recorded. If not available through routine clinical care, blood samples were obtained and tested for CD4 cell count and HIV RNA using the Boston Medical Center Clinical Laboratory. Subjects were followed every 6 months for up to seven observations and follow-up ended in July 2001.

Outcome variables

Uptake of antiretroviral therapy was defined as currently being on antiretroviral therapy at the time of the first interview. Patients reported the names of their antiretroviral medications, as well as the number of doses and the total number of pills prescribed daily. We defined 30-day self-reported adherence as a dichotomous variable and calculated it as the ratio of pills taken over pills prescribed and dichotomized the variable where patients less than 95% adherent were considered non-adherent (Paterson *et al.* 2000). Measurement of HIV viral load was performed using branched-chain DNA (bDNA) techniques (Pachl *et al.* 1995). The lower threshold for detection at the time of the study was 50 copies/ml. There is a variability of 0.3 log (three- to fivefold) in the assay itself, meaning that the difference between sequential values must exceed this difference to be considered clinically significant in an individual patient (Carpenter *et al.* 2000). We defined HIV viral load suppression as achieving HIV RNA of less than 500 copies/ml.

Primary independent variable

We defined substance abuse treatment services as any of the following in the past 6 months: at least 12 weeks in a half-way house or residential facility; at least 12 visits to a substance abuse counselor or mental health professional; day treatment for at least 30 days; or participation in any methadone maintenance program (Brands *et al.* 2002; Farre *et al.* 2002).

Other independent variables

Other specific variables assessed included: age; gender; depressive symptoms as measured by the 20-item Centers for Epidemiologic Studies Depression Scale (CES-D) where ≥ 16 denotes depressive symptoms (Andresen *et al.* 1994); self-reported use of alcohol or drugs in the previous 30 days; severity of alcohol dependence as measured by the Alcohol Dependence Scale (ADS) (Ross *et al.* 1990); homelessness, which was defined as having spent at least 1 night either on the street or in a shelter in the 6 months prior to the interview (Kertesz *et al.* 2003); number of doses of antiretroviral medication per day; CD4 cell count and social support. We measured social support by the Perceived Social Support instrument that has a subscale that assesses general support from friends (Procidano & Heller 1983).

Analysis

We used multivariable logistic regression to examine the association of substance abuse treatment on uptake of antiretroviral therapy at baseline and we adjusted for potential confounding factors by including in the model gender, race, severity of alcohol dependence, homelessness and recent drug injection variables. Multivariable longitudinal logistic regression models were constructed to examine the association of substance abuse treatment on adherence and HIV-RNA (viral load) suppression over time. Because serial measures on the same individuals were considered for the medication adherence and viral load suppression analyses, generalized estimating equations were used to adjust for correlation between these measures over time using a working independence correlation matrix (Liang & Zeger 1986; Zeger & Liang 1986). The substance abuse treatment variable was dichotomous. Two separate models were fitted for 30-day adherence and HIV viral load suppression outcomes. For these models we adjusted for age, gender, depressive symptoms, perceived social support from friends, use of alcohol and drugs in the previous 30 days and doses of antiretroviral medication per day. In the medication adherence and HIV viral load suppression multivariable models, all the predictor variables except for gender and age were allowed to vary with time. All analyses were carried out using SAS.

RESULTS

The 349 subjects had the following baseline characteristics: 79% were men; two-thirds were ethnic minorities; mean age was 41 years; and 29% were homeless. The most common HIV risk factor was injection drug use (59%), with men having sex with men and heterosexual sex each stated by 20% of the non-injecting subjects. Of the injection drug users, 23% had injected drugs in the previous 6 months. In the past 30 days, 24% of subjects reported using alcohol and heroin or cocaine, 18% used alcohol alone and 5% used heroin or cocaine alone; 12% were enrolled in a methadone maintenance program. The average daily alcohol consumption of those drinking in the past 30 days was 6.4 drinks. Over one-third (118/349) of the subjects were engaged in substance abuse treatment at the initial observation. Among the subjects who were not engaged in substance abuse treatment at the initial observation, 70 subjects entered substance abuse treatment during the study period. In terms of primary medical care, 92% (323/349) of the subjects saw a physician two or more times in the preceding 6 months. In this research study, the subjects were followed every 6 months for up to seven occasions and the median number of observations per subject was three. The distribution of interviews (observations) conducted per subject was as follows: 111 subjects had one interview; 40 had two interviews; 48 had three interviews; 44 had four interviews; 39 had five interviews; 47 had six interviews and 20 completed seven interviews. Because study subjects were recruited over a 4-year period, and all follow-ups ceased soon after the end of that recruitment, time of recruitment was the major factor affecting the number of follow-up observations in this study (Ehrenstein *et al.* 2004).

The characteristics of the 205 participants receiving antiretroviral therapy are presented in Table 1. Factors associated with HAART receipt in the bivariate analysis were: not being homeless (22.4% versus 38.2%, $P = 0.001$); and having lower ADS scores (6.9 versus 10.0, $P = 0.006$). In the multivariable model, the following factors were independently associated with receiving antiretroviral therapy (Adjusted OR; 95% CI): use of substance abuse treatment services versus none (1.70; 1.03–2.83) and not being homeless (2.27; 1.34–3.79).

The participant characteristics for 30-day adherence are presented in Table 2. In the multivariate adherence model (Table 3), we found that substance abuse treatment was not associated with 30-day adherence. Depressive symptoms (0.48; 0.32–0.78) and use of drugs or alcohol in the previous 30 days (0.17; 0.11–0.28) were negatively associated with adherence. Factors positively associated with HIV viral load suppression included higher average number of daily doses of antiretroviral medication (1.29; 1.15–1.45) and older age (1.04; 1.00–1.07), whereas use of alcohol and drugs in the previous 30 days was negatively associated with HIV viral load suppression (0.53; 0.35–0.82) (Table 3).

A subanalysis was undertaken to explore whether the absence of a substance abuse treatment effect on adherence was because treatment did not impact on alcohol and drug use. This hypothesis was not borne out and we found that participants in substance abuse treatment had

Table 1 Bivariate and multivariate associations of characteristics of HIV-infected subjects with alcohol problems receiving and not receiving antiretroviral therapy.

Characteristics	*Yes*	*No*		
n	*205*	*144*	*P-value*	*AOR† (95% CI)*
Mean age in years (SD)	41.0 (7.4)	40.0 (7.2)	0.23	
Female (%)	41 (20.0)	32 (22.2)	0.62	
Ethnicity (%)				
Black	84 (41.0)	70 (48.6)	0.25	
White	75 (36.6)	41 (28.5)		
Other	46 (22.4)	33 (22.9)		
Homeless (%)‡	46 (22.4)	55 (38.2)	0.001	2.27 (1.34–3.79)
Jail (%)[Y]	63 (30.7)	39 (27.1)	0.46	
Injected drugs (%)‡	44 (21.5)	37 (25.7)	0.36	
Mean ADS score¶	6.9 (9.8)	10.0 (10.4)	0.006	
SA treatment (%)*	76 (37.1)	42 (29.2)	0.12	1.70 (1.03–2.83)
Mean CD4 cell count (SD)	402 (246)	399 (319)	0.95	
Mean HIV viral load ($\log_{10}$) (SD)	2.0 (1.9)	3.8 (1.5)	0.0001	

†Adjusted odds ratios (AOR) and 95% confidence intervals (95% CI) derived from a multivariable logistic regression model controlling for gender, race, severity of alcohol dependence, and recent drug injection; 337 observations used in this model. ‡In the past 6 months. §Adjusted odds ratio and 95% confidence interval of not being homeless versus being homeless. ¶Alcohol dependence scale (higher score indicates more severe dependence). *Substance abuse treatment: having at least 12 weeks in a half-way house or residential facility; 12 visits to a substance abuse counselor or mental health professional; day treatment for at least 30 days; or participation in any methadone maintenance in the previous 6 months.

Table 2 Bivariate association of subject characteristics and 30-day adherence.

	30-Day adherence		
Characteristics	*Yes (n = 146)*	*No (n = 48)*	*P-value*
Mean age in years (SD)	41.4 (7.3)	39.8 (8.1)	0.22
Female (%)	33 (22.6)	7 (14.6)	0.23
Ethnicity (%)			
Black	63 (43.2)	16 (33.3)	0.48
White	53 (36.3)	20 (41.7)	
Other†	30 (20.6)	12 (25.0)	
Homeless (%).	28 (19.2)	14 (29.2)	0.14
Jail (%)†	45 (30.8)	13 (27.1)	0.62
Injected drugs (%)†	29 (19.9)	13 (27.1)	0.29
Depressive symptoms (%)‡	86 (58.9)	35 (72.9)	0.08
Alcohol and drug use (%)*	52 (35.6)	35 (72.9)	0.0001
Social support scores§			
Friends	9.8 (4.0)	9.0 (4.2)	0.19
SA treatment (%)**	54 (37.2)	18 (36.7)	0.95
Mean doses per day (SD)	5.1 (1.6)	4.6 (1.5)	0.05
Mean CD4 cell count (SD)	414 (254)	375 (216)	0.33
Mean HIV viral load (log_{10}) (SD)	1.8 (1.8)	2.7 (1.9)	0.003

†In the past 6 months. ‡Proportion with a CESD score > 16 indicating depressive symptoms. *In the past 30 days. §Perceived Social Support scale—friend subscale ranges from 1 to 14. **Substance abuse treatment: having at least 12 weeks in a half-way house or residential facility; 12 visits to a substance abuse counselor or mental health professional; day treatment for at least 30 days; or participation in any methadone maintenance in the previous 6 months.

Table 3 Multivariable logistic regression models for the factors associated with 30-day adherence and HIV viral load suppression†

	Adjusted OR (95% CI)	
Factor	*30-day adherence*	*HIV viral load suppression*
SA treatment*	0.72 (0.46–1.12)	0.79 (0.51–1.24)
Depressive‡	0.48 (0.32–0.78)	0.71 (0.46–1.10)
Drugs and alcohol**	0.17 (0.11–0.28)	0.53 (0.35–0.82)
Social support§	1.02 (0.97–1.09)	0.99 (0.93–1.04)
Daily dose¶	0.96 (0.83–1.12)	1.29 (1.15–1.45)
Age	1.02 (0.99–1.06)	1.04 (1.00–1.07)
Male	0.71 (0.37–1.28)	0.62 (0.33–1.16)

†Using generalized estimating equations and controlling for time; 645 observations used in the adherence model and 669 observations used in the HIV viral load model. *SA treatment: having at least 12 weeks in a half-way house or residential facility; 12 visits to a substance abuse counselor or mental health professional; day treatment for at least 30 days; or participation in any methadone maintenance in the previous 6 months. ‡Having a CESD score >= 16 indicating depressive symptoms versus < 16. **Use in the past 30 days. §Perceived Social Support–friend subscale ranges from 1 to 14. ¶Average number of daily doses of antiretroviral medication.

significantly lower odds of using alcohol and drugs in the previous 30 days (adjusted OR; 95% CI = 0.51; 0.33–0.78). Thus, despite the effectiveness of substance abuse treatment on substance use, this exposure did not yield a medication adherence benefit.

DISCUSSION

In this cohort of HIV-infected people with alcohol problems, engagement in substance abuse treatment services and not being homeless were associated with current receipt of antiretroviral therapy. We did not find, however, that engagement of substance abuse treatment services over time had a significant effect on adherence or HIV viral suppression. Consistent with previous literature, the use of drugs or alcohol in the previous 30 days and the presence of depressive symptoms were each negatively associated with self-reported HAART adherence.

Numerous studies have also found that active drug use and not being engaged in addiction treatment was associated with a lower likelihood of receiving antiretroviral treatment (Strathdee *et al.* 1998; Bassetti *et al.* 1999; Moatti *et al.* 2000; Celentano *et al.* 2001; Lucas *et al.*

2001). Conversely, consistent participation in methadone maintenance therapy among HIV-infected opiate users was associated with a higher probability of antiretroviral use (Celentano *et al.* 2001; Turner *et al.* 2001). The association of substance abuse treatment and increased HAART uptake may be due to the willingness of physicians to prescribe HAART to patients whom they believe may be more adherent to medications, or it may be that patients engaged in substance abuse treatment may be more willing to accept HAART. Other studies have also found that homelessness and unstable housing are associated with less access to HAART (Bangsberg *et al.* 1997; Friedland & Williams 1999; Bamberger *et al.* 2000).

Although active drug and alcohol use has been cited as a barrier to achieving optimal adherence to antiretrovirals and HIV viral load suppression (Palepu *et al.* 2003), there are few studies that have addressed the effect of addiction treatment on adherence and virological outcomes. In a longitudinal study of HIV-infected patients attending an urban clinic, Lucas *et al.* found that those who switched from non-use to substance use had worse HAART uptake and adherence, less frequent HIV viral load suppression and lower CD4 cell increases compared to those who remained free of substance use. Conversely, switching from substance use to non-use was strongly associated with improvements in antiretroviral therapy use and adherence, and HIV-1 treatment outcomes, compared to persisting with substance use (Lucas *et al.* 2002). One study found that buprenorphine drug maintenance treatment increased adherence to antiretroviral therapy among HIV-infected drug users (Moatti *et al.* 2000). Among participants in a methadone maintenance program, ongoing illicit injection drug use was associated with non-adherence to antiretroviral therapy but it was not associated with undetectable HIV viral load (Stein *et al.* 2000). Turner *et al.* found recently that among drug-using men enrolled in the New York State Medicaid program, regular drug treatment (defined as at least 6 months duration) was positively associated with a pharmacy-based measure of adherence, but this finding was not observed for women drug users (Turner *et al.* 2003).

In this longitudinal cohort, engagement in substance abuse treatment did not translate to superior HAART adherence or HIV viral suppression. To explain this finding, we tested the hypothesis that the substance abuse treatment services were ineffective at reducing substance use. We found that substance abuse treatment was associated with lower odds of use of alcohol and drugs in the previous 30 days. A possible explanation for not observing an association between substance abuse treatment and adherence may be the lack of integration of HIV care and addiction treatment. A recent study found that almost half of a cohort of HIV-infected drug users, identified through New York State Medicaid files to be on antiretroviral therapy, did not have regular HIV viral load testing. The authors concluded that drug users with HIV-focused care or with regular drug treatment were more likely to have regular HIV viral load testing (Laine *et al.* 2002). The implications of their work as well as our results are that primary care including HIV-focused care and substance abuse treatment may need to be better integrated in order for patients to derive the full benefits of HAART, given that both HIV and addiction are chronic diseases.

Our study has limitations. Although our definition of substance abuse treatment services is not as stringent as that used by Laine *et al.* (2001), we think it has face validity. In our cohort of patients, 36% were categorized as receiving substance abuse treatment services at a reasonable level of exposure, although higher levels of exposure to substance abuse treatment may be required to demonstrate an effect. We were unable to distinguish between current and recent (i.e. within the past 6 months) substance abuse treatment. Our adherence measure is by self-report and tends to over-estimate adherence compared to medication event monitoring systems. However, both measures are strongly correlated and a strong relationship exists between self-reported adherence and HIV-RNA among drug users (Arnsten *et al.* 2001). Finally, we did not have nadir CD4 cell counts values to include in our analyses.

In summary, we found that being engaged in substance abuse treatment is associated with HIV-related medical care among HIV-infected people with alcohol problems. Specifically, substance abuse treatment is associated with current receipt of antiretroviral therapy. However, engagement of substance abuse treatment services over time did not have a significant effect on adherence to antiretroviral medications. Uptake of HAART may be enhanced by effective interactions between the patient and the medical system. Improvements in adherence to HAART are not as yet evident by exposure to substance abuse treatment. Nonetheless, efforts to maximize the effect of substance abuse treatment on adherence to antiretrovirals and HIV treatment outcomes among HIV-infected people with alcohol and drug problems merit further examination in clinical trials (Stone 2001; Ammassari *et al.* 2002; Bartlett 2002; Ickovics & Meade 2002; Tuldra & Wu 2002). In the meantime, substance abuse treatment programs may provide an opportunity to address openly issues of HIV care, including enhancing adherence to HAART.

ACKNOWLEDGEMENTS

The authors appreciate the contributions of the clinical staff of the HIV Diagnostic Evaluation Unit at Boston

Medical Center, including Colleen LaBelle RN and Jennifer Doyle, as well as assistance provided by researchers on the project: Jacqueline Savetsky MPH, Jessica Garcia, Laura Weinstein, Sharon Phillips, Vera Doljanskaia MPH, Kathleen Doyle and Katie Schoolwerth. The authors also appreciate the contributions of Howard Libman MD and the clinical staff at the Beth Israel Deaconess Medical Center. Support for this study came from the following grants from the National Institute on Alcohol Abuse and Alcoholism (NIAAA) of the NIH: RO1-AA13766 (Clinical Impact of HCV and Alcohol in HIV-Infected People); RO1-AA11785 (Medication Adherence in Alcohol Abusing HIV Patients); RO1-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, USPHS grant M01 RR00533. Dr Anita Palepu is supported by the Canadian Institute for Health Research through a New Investigator Award.

REFERENCES

Ammassari, A., Trotta, M. P., Murri, R., Castelli, F., Narciso, P., Noto, P., Vecchiet, J., D'Arminio Monforte, A., Wu, A. W. & Antinori, A. (2002) Correlates and predictors of adherence to highly active antiretroviral therapy: overview of published literature. *Journal of Acquired Immune Deficiency Syndromes*, **31**, S123–S127.

Andresen, E. M., Malmgren, J. A., Carter, W. B. & Patrick, D. L. (1994) Screening for depression in well older adults—evaluation of a short-form of the CES-D. *American Journal of Preventive Medicine*, **10**, 77–84.

Arnsten, J. H., Demas, P. A., Farzadegan, H., Grant, R. W., Gourevitch, M. N., Chang, C. J., Buono, D., Eckholdt, H., Howard, A. A. & Schoenbaum, E. E. (2001) Antiretroviral therapy adherence and viral suppression in HIV-infected drug users: comparison of self-report and electronic monitoring. *Clinical Infectious Diseases*, **33**, 1417–1423.

Bamberger, J. D., Unick, J., Klein, P., Fraser, M., Chesney, M. & Katz, M. H. (2000) Helping the urban poor stay with antiretroviral HIV drug therapy. *American Journal of Public Health*, **90**, 699–701.

Bangsberg, D., Tulsky, J. P., Hecht, F. M. & Moss, A. R. (1997) Protease inhibitors in the homeless. *Journal of the American Medical Association*, **278**, 63–65.

Bartlett, J. A. (2002) Addressing the challenges of adherence. *Journal of Acquired Immune Deficiency Syndromes*, **29**, S2–S10.

Bassetti, S., Battegay, M., Furrer, H., Rickenbach, M., Flepp, M., Kaiser, L., Telenti, A., Vernazza, P. L., Bernasconi, E. & Sudre, P. (1999) Why is highly active antiretroviral therapy (HAART) not prescribed or discontinued? Swiss HIV Cohort Study. *Journal of Acquired Immune Deficiency Syndromes*, **21**, 114–119.

Brands, B., Blake, J. & Marsh, D. (2002) Changing patient characteristics with increased methadone maintenance availability. *Drug and Alcohol Dependence*, **66**, 11–20.

Buchsbaum, D. G., Buchanan, R. G., Centor, R. M., Schnoll, S. H. & Lawton, M. J. (1991) Screening for alcohol-abuse using cage scores and likelihood ratios. *Annals of Internal Medicine*, **115**, 774–777.

Carpenter, C. C., Cooper, D. A., Fischl, M. A., Gatell, J. M., Gazzard, B. G., Hammer, S. M., Hirsch, M. S., Jacobsen, D. M., Katzenstein, D. A., Montaner, J. S., Richman, D. D., Saag, M. S., Schechter, M., Schooley, R. T., Thompson, M. A., Vella, S., Yeni, P. G. & Volberding, P. A. (2000) Antiretroviral therapy in adults: updated recommendations of the International AIDS Society—USA Panel. *Journal of the American Medical Association*, **283**, 381–390.

Carrieri, M. P., Moatti, J. P., Vlahov, D., Obadia, Y., Reynaud-Maurupt, C. & Chesney, M. (1999) Access to antiretroviral treatment among French HIV infected injection drug users: the influence of continued drug use. MANIF 2000 Study Group. *Journal of Epidemiology and Community Health*, **53**, 4–8.

Celentano, D. D., Galai, N., Sethi, A. K., Shah, N. G., Strathdee, S. A., Vlahov, D. & Gallant, J. E. (2001) Time to initiating highly active antiretroviral therapy among HIV-infected injection drug users. *AIDS*, **15**, 1707–1715.

Celentano, D. D., Vlahov, D., Cohn, S., Shadle, V. M., Obasanjo, O. & Moore, R. D. (1998) Self-reported antiretroviral therapy in injection drug users. *Journal of the American Medical Association*, **280**, 544–546.

Ehrenstein, V., Horton, N. J. & Samet, J. H. (2004) Inconsistent condom use among HIV-infected patients with alcohol problems. *Drug and Alcohol Dependence*, **73**, 159–166.

Ewing, J. A. (1984) Detecting alcoholism. The CAGE questionnaire. *Journal of the American Medical Association*, **252**, 1905–1907.

Farre, M., Mas, A., Torrens, M., Moreno, V. & Cami, J. (2002) Retention rate and illicit opioid use during methadone maintenance interventions: a meta-analysis. *Drug and Alcohol Dependence*, **65**, 283–290.

Fiellin, D. A., Reid, C. & O'Connor, P. G. (2000) Screening for alcohol problems in primary care—a systematic review. *Archives of Internatal Medicine*, **160**, 1977–1989.

Folstein, M. F., Folstein, S. E. & McHugh, P. R. (1975) 'Mini-mental state'. A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, **12**, 189–198.

Friedland, G. H. & Williams, A. (1999) Attaining higher goals in HIV treatment: the central importance of adherence. *AIDS*, **13**, S61–S72.

Galvan, F. H., Bing, E. G., Fleishman, J. A., London, A. S., Caetano, R., Burnam, M. A., Longshore, D., Morton, S. C., Orlando, M. & Shapiro, M. (2002) 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. *Journal of Studies on Alcohol*, **63**, 179–186.

Hogg, R. S., Yip, B., Kully, C., Craib, K. J., O'Shaughnessy, M. V., Schechter, M. T. & Montaner, J. S. (1999) Improved survival among HIV-infected patients after initiation of triple-drug antiretroviral regimens. *Canadian Medical Association Journal*, **160**, 659–665.

Ickovics, J. R. & Meade, C. S. (2002) Adherence to HAART among patients with HIV: breakthroughs and barriers. *AIDS Care*, **14**, 309–318.

Kertesz, S. G., Horton, N. J., Friedmann, P. D., Saitz, R. & Samet, J. H. (2003) Slowing the revolving door: stabilization programs reduce homeless people' substance use after detoxification. *Journal of Substance Abuse Treatment*, **24**, 197–207.

Laine, C., Hauck, W. W., Gourevitch, M. N., Rothman, J., Cohen, C. & Turner, B. J. (2001) Regular outpatient medical and drug abuse care and subsequent hospitalization of people who use

illicit drugs. *Journal of the American Medical Association*, 285, 2355–2362.

Laine, C., Zhang, D. Z., Hauck, W. W. & Turner, B. J. (2002) HIV-1 RNA viral load monitoring in HIV-infected drug users on antiretroviral therapy—relationship with outpatient care patterns. *Journal of Acquired Immune Deficiency Syndromes*, 29, 270–274.

Lefevre, F., O'Leary, B., Moran, M., Mossar, M., Yarnold, P. R., Martin, G. J. & Glassroth, J. (1995) Alcohol consumption among HIV-infected patients. *Journal of General Internal Medicine*, 10, 458–460.

Liang, K. Y. & Zeger, S. Y. (1986) Longitudinal data analysis using generalized linear models. *Biometrika*, 73, 13–22.

Lucas, G. M., Cheever, L. W., Chaisson, R. E. & Moore, R. D. (2001) Detrimental effects of continued illicit drug use on the treatment of HIV-1 infection. *Journal of Acquired Immune Deficiency Syndromes*, 27, 251–259.

Lucas, G. M., Gebo, K. A., Chaisson, R. E. & Moore, R. D. (2002) Longitudinal assessment of the effects of drug and alcohol abuse on HIV-1 treatment outcomes in an urban clinic. *AIDS*, 16, 767–774.

Moatti, J. P., Carrieri, M. P., Spire, B., Gastaut, J. A., Cassuto, J. P. & Moreau, J. (2000) Adherence to HAART in French HIV-infected injecting drug users: the contribution of buprenorphine drug maintenance treatment. The Manif 2000 study group. *AIDS*, 14, 151–155.

Mocroft, A., Madge, S., Johnson, A. M., Lazzarin, A., Clumeck, N., Goebel, F. D., Viard, J. P., Gatell, J., Blaxhult, A. & Lundgren, J. D. (1999) A comparison of exposure groups in the EuroSIDA study: starting highly active antiretroviral therapy (HAART), response to HAART, and survival. *Journal of Acquired Immune Deficiency Syndromes*, 22, 369–378.

Moreno, A., Perez-Elias, M. J., Casado, J. L., Munoz, V., Antela, A., Dronda, F., Navas, E. & Moreno, S. (2001) Long-term outcomes of protease inhibitor-based therapy in antiretroviral treatment-naive HIV-infected injection drug users on methadone maintenance programmes. *AIDS*, 15, 1068–1070.

Pachl, C., Todd, J. A., Kern, D. G., Sheridan, P. J., Fong, S. J., Stempien, M., Hoo, B., Besemer, D., Yeghiazarian, T. & Irvine, B. (1995) Rapid and precise quantification of HIV-1 RNA in plasma using a branched DNA signal amplification assay. *Journal of Acquired Immune Deficiency Syndromes*, 8, 446–454.

Palella, F. J. Jr, Delaney, K. M., Moorman, A. C., Loveless, M. O., Fuhrer, J., Satten, G. A., Aschman, D. J. & Holmberg, S. D. (1998) Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators. *New England Journal of Medicine*, 338, 853–860.

Palepu, A., Tyndall, M., Yip, B., O'Shaughnessy, M. V., Hogg, R. S. & Montaner, J. S. G. (2003) Impaired virologic response to highly active antiretroviral therapy associated with ongoing injection drug use. *Journal of Acquired Immune Deficiency Syndromes*, 32, 522–526.

Paterson, D. L., Swindells, S., Mohr, J., Brester, M., Vergis, E. N., Squier, C., Wagener, M. M. & Singh, N. (2000) Adherence to protease inhibitor therapy and outcomes in patients with HIV infection. *Annals of International Medicine*, 133, 21–30.

Petry, N. M. (1999) Alcohol use in HIV patients: what we don't know may hurt us. *International Journal of STD and AIDS*, 10, 561–570.

Procidano, M. E. & Heller, K. (1983) Measures of perceived social support from friends and from family: three validation studies. *American Journal of Community Psychology*, 11, 1–24.

Ross, H. E., Gavin, D. R. & Skinner, H. A. (1990) Diagnostic validity of the MAST and the alcohol dependence scale in the assessment of DSM-III alcohol disorders. *Journal of Studies on Alcohol*, 51, 506–513.

Sambamoorthi, U., Warner, L. A., Crystal, S. & Walkup, J. (2000) Drug abuse, methadone treatment, and health services use among injection drug users with AIDS. *Drug and Alcohol Dependence*, 60, 77–89.

Samet, J. H., Horton, N. J., Dukes, K., Meli, S., Tripp, T., Labelle, C. & Freedberg, K. A. (2002) Randomized controlled trial of a multi-dimensional intervention to enhance adherence to antiretroviral therapy in HIV-infected patients with a history of alcohol problems [abstract]. *XIV International Aids Conference*, Barcelona.

Samet, J. H., Libman, H., LaBelle, C., Steger, K., Lewis, R., Craven, D. E. & Freedberg, K. A. (1995) A model clinic for the initial evaluation and establishment of primary care for people infected with human immunodeficiency virus. *Archives of Internal Medicine*, 155, 1629–1633.

Samet, J. H., Phillips, S. J., Horton, N. J., Traphagen, E. T. & Freedberg, K. A. (2004) Detecting alcohol problems in HIV-infected patients: use of the CAGE questionnaire. *AIDS Research and Human Retroviruses*, in press.

Stein, M. D., Rich, J. D., Maksad, J., Chen, M. H., Hu, P., Sobota, M. & Clarke, J. (2000) Adherence to antiretroviral therapy among HIV-infected methadone patients: effect of ongoing illicit drug use. *American Journal of Drug and Alcohol Abuse*, 26, 195–205.

Stone, V. E. (2001) Strategies for optimizing adherence to highly active antiretroviral therapy: lessons from research and clinical practice. *Clinical Infectious Diseases*, 33, 865–872.

Strathdee, S. A., Palepu, A., Cornelisse, P. G., Yip, B., O'Shaughnessy, M. V., Montaner, J. S., Schechter, M. T. & Hogg, R. S. (1998) Barriers to use of free antiretroviral therapy in injection drug users. *Journal of the American Medical Association*, 280, 547–549.

Tuldra, A. & Wu, A. W. (2002) Interventions to improve adherence to antiretroviral therapy. *Journal of Acquired Immune Deficiency Syndromes*, 31, S154–S157.

Turner, B. J., Laine, C., Cosler, L. & Hauck, W. W. (2003) Relationship of gender, depression, and health care delivery with antiretroviral adherence in HIV-infected drug users. *Journal of General Internal Medicine*, 18, 248–257.

Turner, B. J., Zhang, D., Laine, C., Pomerantz, R. J., Cosler, L. & Hauck, W. W. (2001) Association of provider and patient characteristics with HIV-infected women's antiretroviral therapy regimen. *Journal of Acquired Immune Deficiency Syndromes*, 27, 20–29.

Zeger, S. L. & Liang, K. Y. (1986) Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*, 42, 121–130.

Faculty Self-reported Experience with Racial and Ethnic Discrimination in Academic Medicine

Neeraja B. Peterson, MD, MSc, Robert H. Friedman, MD, Arlene S. Ash, PhD, Shakira Franco, MS, Phyllis L. Carr, MD

BACKGROUND: **Despite the need to recruit and retain minority faculty in academic medicine, little is known about the experiences of minority faculty, in particular their self-reported experience of racial and ethnic discrimination at their institutions.**

OBJECTIVE: **To determine the frequency of self-reported experience of racial/ethnic discrimination among faculty of U.S. medical schools, as well as associations with outcomes, such as career satisfaction, academic rank, and number of peer-reviewed publications.**

DESIGN: **A 177-item self-administered mailed survey of U.S. medical school faculty.**

SETTING: **Twenty-four randomly selected medical schools in the contiguous United States.**

PARTICIPANTS: **A random sample of 1,979 full-time faculty, stratified by medical school, specialty, graduation cohort, and gender.**

MEASUREMENTS: **Frequency of self-reported experiences of racial/ethnic bias and discrimination.**

RESULTS: **The response rate was 60%. Of 1,833 faculty eligible, 82% were non-Hispanic white, 10% underrepresented minority (URM), and 8% nonunderrepresented minority (NURM). URM and NURM faculty were substantially more likely than majority faculty to perceive racial/ethnic bias in their academic environment (odds ratio [OR], 5.4; $P < .01$ and OR, 2.6; $P < .01$, respectively). Nearly half (48%) of URM and 26% of NURM reported experiencing racial/ethnic discrimination by a superior or colleague. Faculty with such reported experiences had lower career satisfaction scores than other faculty ($P < .01$). However, they received comparable salaries, published comparable numbers of papers, and were similarly likely to have attained senior rank (full or associate professor).**

CONCLUSIONS: **Many minority faculty report experiencing racial/ethnic bias in academic medicine and have lower career satisfaction than other faculty. Despite this, minority faculty who reported experiencing racial/ethnic discrimination achieved academic productivity similar to that of other faculty.**

KEY WORDS: **schools, medical; minority groups; faculty, medical; prejudice; job satisfaction.**
J GEN INTERN MED 2004;19:259–265.

Received from the Center for Health Services Research (NBP), Division of General Internal Medicine, Department of Medicine, Vanderbilt University School of Medicine, Nashville, Tenn; and Section of General Internal Medicine (RHF, ASA, SF, PLC), Boston University School of Medicine, Boston, Mass.

This work was presented in part at the National Meeting of the Society of General Internal Medicine, May 2001, San Diego, Calif.

Address correspondence and reprint requests to Dr. Peterson: Center for Health Services Research, Suite 6000, Medical Center East, Nashville, TN 37232 (e-mail: neeraja.peterson@vanderbilt.edu).

Minority faculty make up approximately 17% of full-time faculty in U.S. medical schools; just 4% are from underrepresented minority groups.[1] The underrepresented minorities (URM) are defined as African Americans, Mexican Americans, mainland Puerto Ricans, and American Indians. Each URM group is substantially less prevalent in medicine than in the general population.[2] In comparison, approximately 30% of the U.S. population classify themselves as nonwhite and about 22% are from URM groups.[3] Recruiting and retaining minority faculty in academic medicine is important. Yet, little is known about the faculty experience of minorities, especially with regard to racial and ethnic discrimination, and how such experience affects their career satisfaction and academic success. Our study examines the frequency of self-reported experiences of racial/ethnic bias among faculty in U.S. medical schools, as well as associations of such experience with career satisfaction, and with academic productivity as evidenced by the number of peer-reviewed publications and academic rank.

METHODS

Study Design

In 1995, we conducted a national mailed survey, described in detail elsewhere,[4] to examine the status of minority, women, and generalist faculty in academic medicine. We used 2-stage sampling to select a sample of U.S. medical school faculty. In the first stage, we randomly selected 24 medical schools. Of the 126 medical schools listed by the Association of American Medical Colleges (AAMC) in 1995, we excluded 6 schools outside the contiguous United States because the AAMC considers them to be significantly different from the mainland schools. To obtain adequate numbers of female and minority faculty from each institution, we also excluded 14 schools that had fewer than 200 total faculty, 50 female faculty, or 10 minority faculty. Our 24 medical schools were randomly selected from the remaining 106 eligible medical schools. The resulting sample of schools was balanced across the AAMC's 4 U.S. regions and between public and private institutions.

In the second stage, we selected full-time, salaried faculty members from the 24 schools using the 1994 AAMC Faculty Roster System. The AAMC listed 17,434 faculty at the 24 schools; 720 faculty were excluded because they were in unique departments not found at other medical schools. Of the remaining 16,714 faculty, 4,156 (25%) were women, 929 (6%) were minority, and 869 (6%) were generalists. For each institution, we employed a $4 \times 3 \times 2$ factorial design for stratification. The factors were: 4 areas of medical specialization (primary care, basic science, medical specialties, and surgical specialties), 3 graduation cohorts

(receiving doctorate degree prior to 1970, between 1970 and 1980, and after 1980), and gender. Within each cell (school × medical specialty × graduation cohort × gender), we sought 6 faculty. The most senior cells (by graduation cohort) were filled first and then backfilled, if necessary, with more junior faculty. To obtain sufficient numbers, the sample was supplemented to include all minority, generalist, and senior women faculty. Due to confidentiality concerns of the AAMC, the mailed surveys were delinked from the sampling frame, making it impossible to separately calculate response rates within sampling strata.

Data Collection and Survey Instrument

We mailed 4,405 surveys to sampled faculty, of which 1,073 were ineligible, either because they had left their institutions (512), were not full time (510), had died (11), had participated in the pilot sample (9), or other reasons (31). Of the eligible 3,332 faculty, nonrespondents received reminder postcards, follow-up telephone calls, and survey remailings as necessary. One hundred forty-six respondents were excluded for one or more of the following reasons: they did not self-identify race/ethnicity (30), did not answer questions about bias (42), rank (95), or department (68), or did not complete most of the questionnaire (7).

The self-administered questionnaire asked 177 questions about faculty demographics, experiences of bias, discrimination, and harassment, professional goals and work situation, current academic environment and rank, academic productivity, faculty compensation, and career satisfaction. Approximately 10% of the survey items related to race-based discrimination, which could have occurred at any time over the faculty member's career. The Boston University School of Medicine Institutional Review Board approved this study.

Definitions of Variables and Outcome Measures

We divided faculty respondents' self-reported race/ethnicity into 3 categories as defined by the AAMC: underrepresented minority (URM; non-Hispanic Black; Mexican American and Puerto Rican Hispanic; Native American or Alaskan Native), nonunderrepresented minority (NURM; Asian and other Hispanic groups), and majority (non-Hispanic white). We coded specialties as follows: primary care (general internal medicine, general pediatrics, family medicine, and geriatrics); medical specialties (internal medicine subspecialties, pediatric subspecialties, neurology, physical medicine, radiology, emergency medicine, anesthesia, and psychiatry); surgical specialties (general surgery and its subspecialties); and basic science. We asked respondents to estimate the number of hours worked during an average professional work week and the amount of time spent in research, patient care, teaching, and administration.

We asked 3 questions to characterize the experience of racial/ethnic bias: 1) "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 to 5 = yes, frequently)?"; 2) "In your professional career, have you ever been left out of opportunities for professional advancement based on race/ethnicity (1 = no, 2 = not to my knowledge, 3 = possibly, 4 = probably, 5 = yes)?"; and 3) "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)?"

Faculty who answered "yes" to question 3 were asked 5 questions to capture the extent and severity of the racial/ethnic discrimination they experienced: 1) "How much of a problem has this been for you (1 = no problem to 5 = major problem)?"; 2) "Have you encountered racial/ethnic remarks (1 = no, 2 = yes)?"; 3) "Have you encountered inadequate recognition of your work (1 = no, 2 = yes)?"; 4) "To what extent have these experiences had a negative effect on your confidence as a professional (1 = not at all to 5 = greatly)?"; and 5) "To what extent have these experiences negatively affected your career advancement (1 = not at all to 5 = greatly)?"

We also asked all faculty about several subjective and objective outcomes. To capture career satisfaction, we used a scale consisting of 4 items: "How satisfied are you with 1) your current work setting, 2) your potential to achieve your professional goals, 3) your overall professional practice and/or research, and 4) the extent to which this practice and/or research has met your expectations?"[5] Each item was measured on a Likert scale ranging from 1 (very dissatisfied) to 5 (very satisfied) (Cronbach's $\alpha = 0.87$). Two other subjective outcome variables were measured from questions with 5-point Likert scales. These were 1) "To what extent do you feel like a welcomed member in your institution?" and 2) "How likely are you to leave academic medicine within 5 years and go into another line of work?"

Career outcome variables included senior rank (associate or full professor), salary, total career publications in refereed journals, and grants funded. Salary was 1995 pretax faculty compensation and included salary and other professional payments, but excluded fringe benefits. Grants funded was the number of grants with the respondent as the principal investigator that had received funding within the previous 2 years. Missing responses to numbers of publications or grants funded were treated as zeros.

Analysis

Frequency distributions, means, and standard deviations of characteristics were used to describe the survey respondents by minority status (URM, NURM, and majority). The distributions of characteristics among majority faculty reflect the factorial sampling design. For example, the design sought approximately even numbers of men and women majority faculty. In contrast, minority respondent characteristics reflect a full-census sampling of all minority faculty at the selected medical schools.

For questions that captured the perceptions or experiences of racial/ethnic bias, we used a 5-point Likert

scale, and scored any response of 3, 4, or 5 as positive. The magnitude of differences among racial/ethnic groups did not significantly change when responses of only 4 or 5 were scored as positive.

We used multivariable analyses to test for relationships between faculty characteristics and perceptions and/or experiences of racial/ethnic bias. The following variables appeared in all models: medical school, specialty (primary care, basic science, and medical and surgical specialties), minority status, gender, seniority (years since first faculty appointment), and seniority squared (to capture the declining influence of additional years on outcomes).

Analyses were performed using SAS statistical software, version 8.2 (SAS Institute Inc., Cary, NC). We used Fisher's exact test to compare the frequency of racial/ethnic discrimination by minority status, and linear regression (PROC GLM) to estimate the effects of having experienced racial/ethnic discrimination on feelings of welcomeness, likelihood of leaving the current institution, career satisfaction, salary, number of publications, and number of grants funded. We used logistic regression (PROC LOGISTIC) to estimate the effect on attainment of senior rank. In additional analyses, we also controlled for number of hours worked per week, percentage of time in research, and percentage of time in clinical work. Because we found few differences compared to the models using the original linear regression covariates described above, the results of the models using only the original list of variables are reported. We tested for multicollinearity between dependent variables in our models with the TOL and VIG options. We tested for interactions between minority status and the experience of racial/ethnic discrimination on all outcomes (career satisfaction, feelings of welcomeness, likelihood of leaving the current institution, attainment of senior rank, salary, number of publications, and number of grants funded). We used mixed-effects regression modeling (PROC MIXED) to address clustering by medical school and compared the results to the findings using PROC GLM.

None of the authors had any potential conflicts of interest. Authors had full access to all of the data in the study, and accept full responsibility for the integrity of the data and the accuracy of the data analysis. The Robert Wood Johnson Foundation funded the study but had no role in its design, conduct, or reporting.

RESULTS

Demographics and Professional Characteristics

Of the 3,332 eligible faculty study subjects, 1,979 returned the survey for a response rate of 60%. Eighty-two percent of respondents identified themselves as non-Hispanic white (majority), 10% as URM (Blacks [8%], Mexican Americans [1%], Puerto Ricans [1%], Native Americans or Alaskan Natives [0.3%]), and 8% as NURM (Asian or Pacific Islanders [7%] and other Hispanic Americans [1%]).

Table 1 shows the demographic and professional characteristics of faculty respondents by the 3 racial/ethnic groups. Majority respondents were on average 2 years older and had been on the faculty for 2 years longer than minority respondents. More URM faculty were male (60%) than NURM (40%). NURM differed from the other 2 groups in how few were born in the United States (21% vs 74% and 88%, respectively) or had English as their primary language (65% vs 85% and 97%). The URM faculty were more likely to be in a medical specialty (45% vs 28% and 24%) and spent more time in clinical activities (39% vs 32% and 32%). They were less likely to be in the basic sciences (12% vs 28% and 25%), to be a full or associate professor (31% vs 41% and 58%), and had fewer career publications (15 vs 22 and 32).

Perceived Bias Attributed to Faculty Race/Ethnicity

Table 2 shows frequency of perceived racial/ethnic bias by minority status. Most (63%) of the URM faculty perceived racial/ethnic bias or obstacles to the career success or satisfaction of faculty in their academic environment compared to 50% of NURM and 29% of majority faculty. In the multivariable analyses, URM faculty had 5.4 times the odds of perceiving racial/ethnic bias in their academic environment than the majority faculty; NURM were also more likely than the majority faculty to perceive these problems (odds ratio [OR], 2.6). In addition, URM faculty and NURM faculty also significantly had more odds than the majority of having reported experiencing racial/ethnic bias in their professional advancement (OR, 12.8 and 6.9, respectively). Nearly half (48%) of URM and 26% of NURM faculty reported personally encountering racial/ethnic discrimination by a superior or colleague compared to 7% of the majority faculty.

Factors Associated with the Perception of Racial/Ethnic Bias

Other faculty characteristics were examined for associations with racial/ethnic bias. The following were significantly associated with reporting personal experiences of racial/ethnic bias: 1) increasing age (OR, 1.5 per 10 years; 95% confidence interval [CI], 1.1 to 2.1); 2) having a primary language other than English (OR, 1.8; 95% CI, 1.1 to 3.0); and 3) increasing number of hours worked (OR, 1.3 per 10 hours/week; 95% CI, 1.1 to 1.5). Medical school characteristics, including which medical school a faculty member was at, whether it was a private or public institution, and its regional location, were not significantly associated with reporting personal experiences of racial/ethnic bias.

Most faculty who reported personal experiences of racial/ethnic discrimination stated instances in which their work was inadequately recognized (78%, 78%, and 63% of URM, NURM, and majority, respectively; $P = .05$) (Table 3). Most also reported personally encountering racial/ethnic remarks ($P = .29$). Smaller numbers of faculty felt that the discrimination had been a major problem for them (32%, 19%, and 28% of URM, NURM, and majority, respectively; $P = .40$) or that it had a major effect on their professional

Table 1. Respondent Demographics and Professional Characteristics by Minority Status*

	Underrepresented Minority (*N* = 185)	Nonunderrepresented Minority (*N* = 141)	Majority (*N* = 1,507)
Demographics			
Mean age, y ± SD	44.1 ± 9.2	44.3 ± 9.5	46.5 ± 9.4
Male, %	60	40	51
Born in U.S., %	74	21	88
English as primary language, %	85	65	97
Professional characteristics			
Years as faculty, mean ± SD	9.7 ± 8.4	10.3 ± 9.0	12.0 ± 9.2
Medical degree (MD), %	79	70	67
Hours worked per week, mean ± SD	59 ± 14	58 ± 13	57 ± 12
Percent time spent in, mean ± SD			
Clinical activities	39 ± 27	32 ± 30	32 ± 29
Research	20 ± 25	34 ± 35	28 ± 29
Teaching	21 ± 15	17 ± 12	21 ± 15
Administration	19 ± 20	16 ± 18	19 ± 18
Specialty, %			
Primary care	26	26	32
Basic science	12	28	25
Medical specialty	45	28	24
Surgical specialty	17	17	18
Rank, %			
Full professor	16	18	30
Associate professor	15	23	28
Assistant professor	62	51	37
Instructor	8	8	4
U.S. region, %			
Northeast	35	40	38
South	21	15	23
Midwest	19	22	19
West	23	22	19
Public institution, %	59	52	51
Salary in thousands, mean ± SD	115 ± 63	101 ± 51	111 ± 58
Total career publications, mean ± SD	15.2 ± 25.8	22.0 ± 30.6	31.5 ± 42.2
Grants funded, mean ± SD†	1.1 ± 3.0	0.9 ± 1.7	1.2 ± 2.1
Career satisfaction score, mean ± SD‡	3.2 ± 0.9	3.4 ± 0.8	3.5 ± 0.9

* *Information is missing on gender for 2, marital status for 15, country of birth for 1, primary language for 1, degree for 38, hours worked per week for 5, U.S. region for 21, institution for 12, and salary for 89.*

† *In the preceding 2 years.*

‡ *From McGlynn's 4-item scale. Each item was measured on a Likert scale of 1 to 5 (1 = very dissatisfied, 5 = very satisfied).*

SD, standard deviation.

confidence (17%, 19%, and 18% of URM, NURM, and majority, respectively; $P = .95$). About a third of URM and NURM faculty who reported experiencing racial/ethnic discrimination felt that the experience had a major effect on their career advancement ($P = .51$).

Career Satisfaction and Personal Experience of Racial/Ethnic Discrimination

Faculty who had personally experienced racial/ethnic bias had lower career satisfaction scores than other faculty (adjusted mean scores, 3.2 vs 3.5, respectively; $P < .01$), and were less likely to feel welcomed at their institution (adjusted mean scores, 3.3 vs 3.9, respectively; $P < .01$) (Table 4). However, faculty who experienced racial/ethnic bias reported themselves as not more likely to leave academic medicine within 5 years (adjusted mean scores, 2.5 vs 2.3, respectively; $P = .17$). There was no evidence of collinearity in the dependent variables of minority status and personal experience of discrimination. There were no significant interactions between minority status and the experience of racial/ethnic discrimination at the $P \leq .05$ level on career satisfaction outcomes. Mixed-effects regression modeling did not alter any study findings.

Associations with Career Outcomes

There were no statistically significant associations between the personal experience of racial/ethnic bias and attainment of career outcomes including senior rank (full or associate professor), salary, number of career publications, or number of grants funded in the previous 2 years (all $P > .1$) (Table 5). We found that URM faculty and NURM faculty were less likely to attain senior rank (OR, 4.4; 95% CI, 2.6 to 7.7 and OR, 2.0; 95% CI, 1.1 to 3.5, for URM and NURM, respectively) after adjustment for self-reported

Table 2. Perception and Experience of Racial/Ethnic Bias by Minority Status*

	Reported Percent	Adjusted OR†	95% CI
Respondents who perceived racial/ethnic bias in the academic environment‡			
URM	63	5.4	3.8 to 7.8
NURM	50	2.6	1.8 to 3.7
Majority	29	1.0	–
Respondents who personally experienced racial/ethnic bias in professional advancement§			
URM	54	12.8	8.7 to 18.7
NURM	36	6.9	4.5 to 10.5
Majority	8	1.0	–
Respondents who personally experienced racial/ethnic discrimination by a superior or colleague‖			
URM	48	12.3	8.4 to 18.2
NURM	26	5.0	3.2 to 7.8
Majority	7	1.0	–

* *Underrepresented minorities (URM): n = 185; nonunderrepresented minorities (NURM): n = 141; majority: n = 1,507.*

† *Adjusted for medical school, specialty, gender, and years since first faculty appointment. All P values <.01.*

‡ *Five-point Likert scale with 1 = no, never; 5 = yes, frequently, and 3 to 5 scored as positive. If only 4 or 5 was scored as positive, "Reported Percent" were 41% for URM, 29% for NURM, and 14% for majority.*

§ *Five-point Likert scale with 1 = no, 2 = not to my knowledge, 3 = possibly, 4 = probably, 5 = yes, and 3 to 5 scored as positive. If only 4 or 5 was scored as positive, "Reported Percent" were 33% for URM, 19% for NURM, and 4% for majority.*

‖ *1 = no, 2 = yes.*

OR, odds ratio; CI, confidence interval.

personal experiences of discrimination, medical school, specialty, minority status, gender, seniority, and seniority squared. Tests for effect modification revealed no significant interactions between minority status and the experience of racial/ethnic discrimination on career outcomes, and there was no evidence of collinearity between the 2 dependent variables. Additionally, mixed-effects regression modeling did not alter any study findings.

DISCUSSION

Little is known about minority faculty's experience with racial and ethnic discrimination in academic medicine. In our study of a national sample of academic faculty, we were able to address both subjective perceptions and objective career outcomes of racial/ethnic discrimination, not just its frequency. We found that substantial numbers of both URM and NURM faculty perceived racial bias in their academic environment, while majority faculty infrequently perceived such bias. Nearly half of URM and over a quarter of NURM faculty reported personal encounters with racial/ethnic discrimination by a superior or a colleague.

Having a primary language other than English was associated with the experience of racial/ethnic bias, independent of minority status; we can speculate that having accented speech may make some faculty have "outsider" status. In addition, older faculty perceived more racial/ethnic bias. This finding may indicate a real improvement in that younger minority faculty are less likely to have a negative experience. However, it may simply reflect that longer careers provide more opportunity to encounter bias.

Previous studies have shown disparities in the promotion of minority faculty. Petersdorf et al. reported that minority faculty with an MD degree in 1989 were promoted to the associate professor level 3 to 7 years later than white

Table 3. Perceptions Among Faculty Reporting a Personal Experience of Racial/Ethnic Discrimination by Minority Status

Effect	Underrepresented Minority, % (N = 89)	Nonunderrepresented Minority, % (N = 37)	Majority, % (N = 104)	P Value
Personally encountered inadequate recognition of work*	78	78	63	.05
Personally encountered racial/ethnic remarks*	79	69	70	.29
Racial/ethnic bias has been a major problem for me†‡	32	19	28	.40
Racial/ethnic bias has had a major effect on the following†§				
Professional confidence	17	19	18	.95
Career advancement	32	33	22	.51

* *1 = no, 2 = yes.*

† *Five-point Likert scale with 1 = no problem, 5 = major problem.*

‡ *Major problem or effect defined as a response of 4 or 5 compared to responses of 1 or 2.*

§ *Five-point Likert scale with 1 = not at all, 5 = greatly.*

Table 4. Career Satisfaction by Personal Experience of Racial/Ethnic Discrimination

Outcome	Personally Experienced Discrimination (*N* = 230)	Did Not Personally Experience Discrimination (*N* = 1,603)	Adjusted *P* Value*
Mean career satisfaction score†	3.2 ± 0.06	3.5 ± 0.02	< .01
Felt like a welcomed member in institution‡	3.3 ± 0.08	3.9 ± 0.03	< .01
Likely to leave academic medicine within 5 years§	2.5 ± 0.11	2.3 ± 0.04	.17

* *Adjusted for medical school, specialty, minority status, gender, and years since first faculty appointment.*
† *From McGlynn's 4-item scale. Each item was measured on a Likert scale of 1 to 5 (1 = very dissatisfied, 5 = very satisfied).*
‡ *Five-point Likert scale with 1 = unwelcome, 5 = fully welcomed.*
§ *Five-point Likert scale with 1 = not at all likely, 5 = very likely.*

faculty.[2] In an earlier study of this faculty sample, we showed that URM faculty were significantly more likely than majority faculty to not hold senior academic rank (OR, 3.4; 95% CI, 1.9 to 6.3 for URM, and OR, 1.6; 95% CI, 0.8 to 2.9 for NURM not holding senior rank, respectively, compared to majority faculty),[6] similar to our results. Similarly, Fang et al. showed that URM faculty at the assistant professor rank and at the associate professor rank were less likely to be promoted when compared to majority faculty.[7] Racial/ethnic discrimination may be the reason for the promotion disparity;[7,8] however, our current study did not find faculty's personal experiences of racial/ethnic bias was associated with attainment of senior rank (full or associate professor) independent of minority status.

Our study showed that faculty who experienced racial/ethnic bias were less likely to feel satisfied with their careers and less likely to feel welcomed in their institutions than those who did not, and the difference was a "medium" effect size.[9] This may explain why URM faculty as a group has been found to be less satisfied with their careers.[10] This lack of satisfaction and belonging was present despite comparable salaries, numbers of publications, and grants. This finding may reflect that minority faculty are able to overcome their negative experiences at their institutions and still achieve high productivity in academic medicine. However, it may also reflect that we did not capture the true experience of all minority faculty because we did have a 40% nonresponse rate to our survey. We also had no way of capturing the experience of minority faculty who had already left academic medicine. To the extent that discrimination contributes to leaving, we may have underrepresented the frequency of racial/ethnic bias, and underestimated its professional impact.

The major limitation of our study is that it is cross-sectional and cannot follow the effects of racial/ethnic discrimination on faculty careers over time. Even though we report associations of racial/ethnic discrimination with several outcomes, we cannot determine cause and effect. For example, we cannot distinguish whether the perception of racial/ethnic bias results in lower job satisfaction or whether lower job satisfaction increases the perception of racial/ethnic bias. Our self-reported questionnaire format is not able to explore the qualitative experience of racial/ethnic discrimination. We examined racial/ethnic discrimination by superiors and colleagues only and did not explore other possible sources of such problems, including patients and hospital staff. We do not know how well our data reflect the current academic environment for minority faculty, as the discrimination that we captured could have occurred at any point in the academic careers of the respondents. Finally, the results that we report are several years old. Since 1995, academic institutions have continued to place increasing importance on minority faculty issues and cultural competence. Thus, the academic environment for minority faculty may have significantly improved since our study was conducted.

Table 5. Career Outcomes by Personal Experience of Racial/Ethnic Discrimination

Outcome	Personally Experienced Discrimination (*N* = 230)	Did Not Personally Experience Discrimination (*N* = 1603)	Adjusted *P* Value*
Senior rank†	OR, 1.1 (95% CI, 0.7 to 1.7)	–	.77
Salary, × $1000 ± SD	107 ± 3.6	112 ± 1.3	.25
Total career publications, *n* ± SD	26.0 ± 2.5	28.9 ± 0.9	.27
Grants funded, *n* ± SD‡	1.3 ± 0.2	1.1 ± 0.1	.15

* *Adjusted for medical school, specialty, minority, gender, and years since first faculty appointment.*
† *Full professor or associate professor.*
‡ *In the preceding 2 years.*
OR, odds ratio; CI, confidence interval.

Our study has several strengths. We determined the frequency of racial/ethnic bias among a large group of medical faculty across all medical school departments using a national database. Because our study was part of a larger study examining the status of faculty in academic medicine, response bias should be less than in a more narrowly focused study of racial bias and discrimination only.

The high frequency of perceived racial/ethnic discrimination among minority faculty is concerning. Understanding the reasons for this and addressing the causes is both a moral and social issue for medical schools and teaching hospitals. In our study, we were not able to show that racial/ethnic discrimination explained the disparities in academic advancement found in other studies and our previous work. Therefore, other explanations for disparities in academic promotion among minority faculty must be pursued.

We thank Anita Palepu, MD, MPH for her assistance and expertise on minority faculty issues. We would also like to acknowledge Mark A. Moskowitz, MD (deceased) for his contributions to the design of this study. This work was supported in part by a grant from the Robert Wood Johnson Foundation.

REFERENCES

1. U.S. Medical School Faculty, 2000. Available at: http://www.aamc.org/findinfo/infores/datarsc/facros/frspubs/usmsf00/start.htm. Accessed July 17, 2001.
2. Petersdorf RG, Turner KS, Nickens HW, Ready T. Minorities in medicine: past, present, and future. Acad Med. 1989;65:663–70.
3. Profile of General Demographic Characteristics, 2000 Census of Population and Housing, United States. Available at: http://www.census.gov/prod/cen2000/dp1/2kh00.pdf. Accessed August 17, 2001.
4. Carr PL, Ash AS, Friedman RH, et al. Faculty perceptions of gender discrimination and sexual harassment in academic medicine. Ann Intern Med. 2000;132:889–96.
5. McGlynn EA. Physician Job Satisfaction. Its Measurement and Use as an Indicator of System Performance. Santa Monica, Calif: RAND; 1989.
6. Palepu A, Carr PL, Friedman RH, Ash AS, Moskowitz MA. Minority faculty and academic rank in medicine. JAMA. 1998;280:767–71.
7. Fang D, Moy E, Colburn L, Hurley J. Racial and ethnic disparities in faculty promotion in academic medicine. JAMA. 2000;284:1085–92.
8. Cohen JJ. Time to shatter the glass ceiling for minority faculty. JAMA. 1998;280:821–2.
9. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd edn. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
10. Palepu A, Carr PL, Friedman RH, Ash AS, Moskowitz MA. Specialty choices, compensation, and career satisfaction of underrepresented minority faculty in academic medicine. Acad Med. 2000;75:157–60.

Risk Adjustment of Medicare Capitation Payments Using the CMS-HCC Model

Gregory C. Pope, M.S., John Kautter, Ph.D., Randall P. Ellis, Ph.D., Arlene S. Ash, Ph.D., John Z. Ayanian, M.D., M.P.P., Lisa I. Iezzoni, M.D., M.Sc., Melvin J. Ingber, Ph.D., Jesse M. Levy, Ph.D., and John Robst, Ph.D.

This article describes the CMS hierarchical condition categories (HCC) model implemented in 2004 to adjust Medicare capitation payments to private health care plans for the health expenditure risk of their enrollees. We explain the model's principles, elements, organization, calibration, and performance. Modifications to reduce plan data reporting burden and adaptations for disabled, institutionalized, newly enrolled, and secondary-payer subpopulations are discussed.

INTRODUCTION

Medicare is one of the world's largest health insurance programs, with annual expenditures exceeding $200 billion. It provides health insurance to nearly 40 million beneficiaries entitled by elderly age, disability, or ESRD. Approximately 11 percent of Medicare beneficiaries are enrolled in private managed care health care plans, with the rest in the traditional FFS program. The 1997 BBA modified the Medicare managed care (MMC) and other capitated programs, collectively called M+C.[1] Medicare pays private plans participating in M+C a monthly capitation rate to provide health care services to enrolled beneficiaries.

Historically, capitation payments to MMC plans were linked to FFS expenditures by geographic area, with payments set at 95 percent of an enrollee's county's adjusted average per capita cost (AAPCC). The AAPCC actuarial rate cells were defined by: age, sex, Medicaid enrollment (indicating poverty), institutional status (for nursing home residents), and working aged status (for beneficiaries with employer-based insurance where Medicare is a secondary payer). Separate county factors were calculated for the aged and non-aged disabled (under 65 years), and at the State-level only (due to small numbers), for ESRD-entitled beneficiaries.

The AAPCC payment methodology explains only about 1-percent of the variation in expenditures for Medicare beneficiaries, and does not pay more for sicker people. Thus, research showed that the managed care program was increasing total Medicare Program expenditures, because its enrollees were healthier than FFS enrollees, and the AAPCC did not account for this favorable selection (Brown et al., 1993; Riley et al., 1996; Mello et al.,

Gregory C. Pope and John Kautter are with RTI International. Randall P. Ellis and Arlene S. Ash are with Boston University. John Z. Ayanian is with Harvard Medical School and Brigham and Women's Hospital. Lisa I. Iezzoni is with Harvard Medical School and Beth Israel Deaconess Medical Center. Melvin J. Ingber, Jesse M. Levy, and John Robst are with the Centers for Medicare & Medicaid Services (CMS). The research in this article was funded by CMS to RTI International under Contract Numbers 500-95-048 and 500-00-0030. The views expressed in this article are those of the authors and do not necessarily reflect the views of RTI International, Boston University, Harvard Medical School, Brigham and Women's Hospital, Beth Israel Deaconess Medical Center, or CMS.

[1] The Medicare Prescription Drug, Improvement and Modernization Act of 2003 (MMA) renames the M+C program Medicare Advantage. However, since this renaming does not officially take place until 2006, we continue to use M+C.

2003). Also, more money was not directed to plans enrolling sicker beneficiaries, or to plans specializing in treating high-cost populations, such as beneficiaries with particular chronic diseases or high levels of functional impairment.

The M+C program fundamentally changed the MMC payment method, including a mandate for health-based Medicare capitation payments by 2000. To support this mandate, the BBA required managed care organizations (MCOs) to report inpatient encounter data (i.e., records for each inpatient admission of a plan's enrollees noting, among other things, the beneficiaries' diagnoses) beginning in 1998. In 2000 CMS, which administers the Medicare Program, implemented the PIP-DCG model as a health-based payment adjuster (Pope et al., 2000a). This model estimates beneficiary health status (expected cost next year) from AAPCC-like demographics and the worst principal inpatient diagnosis (principal reason for inpatient stay) associated with any hospital admission. PIP-DCG-based payments were introduced gradually, with only 10 percent of total Medicare capitation payments adjusted by PIP-DCG factors in 2000. The other 90 percent of payments were still adjusted using a purely demographic (AAPCC-like) model.

The PIP-DCG model was intended as a transition, a feasible way to implement risk adjustment based on the readily available, already audited inpatient diagnostic data. Relying on inpatient diagnoses is the PIP-DCG model's major shortcoming, since only illnesses that result in hospital admissions are counted; MCOs that reduce admis-sions (e.g., through good ambulatory care) can end up with apparently healthier patients and lower payments. Congress's BIPA (2000) addressed the PIP-DCG limitations by requiring the use of ambulatory diagnoses in Medicare risk-adjustment, to be phased in from 2004 to 2007 at 30, 50, 75, and 100 percent of total payments. CMS began collecting encounter data from MCOs for the physician office and hospital outpatient settings (i.e., records of each enrollee visit to these providers with dates, procedures performed, diagnoses, etc.) in October 2000 and April 2001, respectively. However, following complaints from MCOs about the burden of reporting encounter data, CMS suspended data collection in May 2001, ultimately adopting a drastically streamlined data reporting strategy (discussed later).

CMS evaluated several risk-adjustment models that use both ambulatory and inpatient diagnoses, including ACGs (Weiner et al., 1996), the chronic disease and disability payment system (CDPS) (Kronick et al., 2000), clinical risk groups (CRGs) (Hughes et al., 2004), the clinically detailed risk information system for cost (CD-RISC) (Kapur et al., 2003), and DCG/HCCs (Pope et al, 2000b). CMS chose the DCG/HCC model for Medicare risk-adjustment, largely on the basis of transparency, ease of modification, and good clinical coherence. The DCG/HCC model, part of the same DCG family of models as the PIP-DCG, was developed with CMS funding by researchers at RTI International[2] and Boston University, with clinical input from physicians at Harvard Medical School.[3]

Prior to implementing Medicare risk-adjustment in 2004, the DCG/HCC model developers and CMS staff adapted the original model for consistency with CMS' simplified data collection, and for customized fit for Medicare subpopulations. The resulting CMS-HCC model reflects these

2 The early development of the DCG/HCC model was done by Health Economics Research, Inc. while under contract to CMS. However, RTI International acquired Health Economics Research, Inc. in 2002.

3 The original version of the DCG/HCC model is described in Ellis et al. (1996). The DCG/HCC model has been refined as described in Pope et al., 1998 and 2000b.

Medicare-specific adaptations of the DCG/HCC model and provides a comprehensive framework for Medicare risk-adjustment.

This article describes the DCG/HCC and CMS-HCC models. The next section describes the DCG/HCC model, including the principles and elements of its diagnostic classification system and how its performance compares to earlier models. We then describe the modifications to accommodate the simplified data that lead to the CMS-HCC model. The final section describes the CMS-HCC model adaptations for subpopulations.

DCG/HCC MODEL PRINCIPLES

Diagnostic Classification System

The following ten principles guided the creation of the diagnostic classification system.

Principle 1—Diagnostic categories should be clinically meaningful. Each diagnostic category is a set of ICD-9-CM codes (Centers for Disease Control and Prevention, 2004). These codes should all relate to a reasonably well-specified disease or medical condition that defines the category. Conditions must be sufficiently clinically specific to minimize opportunities for gaming or discretionary coding. Clinical meaningfulness improves the face validity of the classification system to clinicians, its interpretability, and its utility for disease management and quality monitoring.

Principle 2—Diagnostic categories should predict medical expenditures. Diagnoses in the same HCC should be reasonably homogeneous with respect to their effect on both current (this year's) and future (next year's) costs. (In this article we present prospective models predicting future costs.)

Principle 3—Diagnostic categories that will affect payments should have adequate sample sizes to permit accurate and stable estimates of expenditures. Diagnostic categories used in establishing payments should have adequate sample sizes in available data sets. Given the extreme skewness of medical expenditure data, the data cannot reliably determine the expected cost of extremely rare diagnostic categories.

Principle 4—In creating an individual's clinical profile, hierarchies should be used to characterize the person's illness level within each disease process, while the effects of unrelated disease processes accumulate. Because each new medical problem adds to an individual's total disease burden, unrelated disease processes should increase predicted costs of care. However, the most severe manifestation of a given disease process principally defines its impact on costs. Therefore, related conditions should be treated hierarchically, with more severe manifestations of a condition dominating (and zeroing out the effect of) less serious ones.

Principle 5—The diagnostic classification should encourage specific coding. Vague diagnostic codes should be grouped with less severe and lower-paying diagnostic categories to provide incentives for more specific diagnostic coding.

Principle 6—The diagnostic classification should not reward coding proliferation. The classification should not measure greater disease burden simply because more ICD-9-CM codes are present. Hence, neither the number of times that a particular code appears, nor the presence of additional, closely related codes that indicate the same condition should increase predicted costs.

Principle 7—Providers should not be penalized for recording additional diagnoses (monotonicity). This principle has

two consequences for modeling: (1) no condition category should carry a negative payment weight, and (2) a condition that is higher-ranked in a disease hierarchy (causing lower-rank diagnoses to be ignored) should have at least as large a payment weight as lower-ranked conditions in the same hierarchy.

Principle 8—The classification system should be internally consistent (transitive). If diagnostic category A is higher-ranked than category B in a disease hierarchy, and category B is higher-ranked than category C, then category A should be higher-ranked than category C. Transitivity improves the internal consistency of the classification system, and ensures that the assignment of diagnostic categories is independent of the order in which hierarchical exclusion rules are applied.

Principle 9—The diagnostic classification should assign all ICD-9-CM codes (exhaustive classification). Since each diagnostic code potentially contains relevant clinical information, the classification should categorize all ICD-9-CM codes.

Principle 10—Discretionary diagnostic categories should be excluded from payment models. Diagnoses that are particularly subject to intentional or unintentional discretionary coding variation or inappropriate coding by health plans/providers, or that are not clinically or empirically credible as cost predictors, should not increase cost predictions. Excluding these diagnoses reduces the sensitivity of the model to coding variation, coding proliferation, gaming, and upcoding.

In designing the diagnostic classification, principles 7 (monotonicity), 8 (transitivity), and 9 (exhaustive classification) were followed absolutely. For example, if the expenditure weights for our models did not originally satisfy monotonicity, we imposed constraints to create models that did. Judgment was used to make tradeoffs among other principles. For example, clinical meaningfulness (principle 1) is often best served by creating a very large number of detailed clinical groupings. But a large number of groupings conflicts with adequate sample sizes for each category (principle 3). Another tradeoff is encouraging specific coding (principle 5) versus predictive power (principle 2). In current coding practice, non-specific codes are common. If these codes are excluded from the classification system, substantial predictive power is sacrificed. Similarly, excluding discretionary codes (principle 10) can also lower predictive power (principle 2). We approached the inherent tradeoffs involved in designing a classification system using empirical evidence on frequencies and predictive power, clinical judgment on relatedness, specificity, and severity of diagnoses, and the judgment of the authors on incentives and likely provider responses to the classification system. The DCG/HCC models balance these competing goals to achieve a feasible health-based payment system.

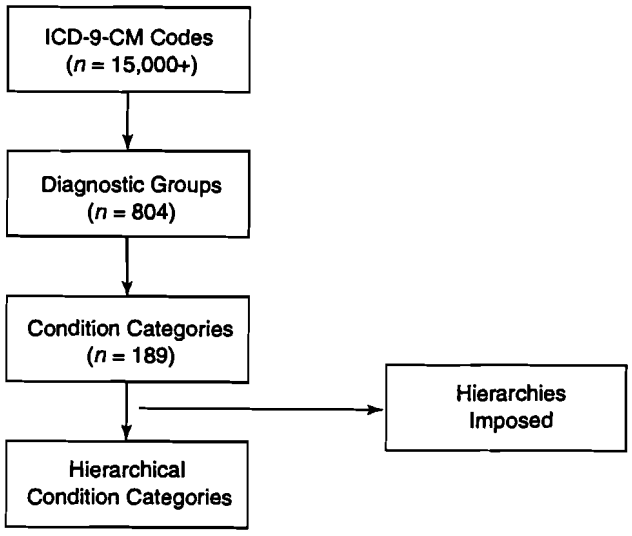
Elements and Organization

As shown in Figure 1, the HCC diagnostic classification system first classifies each of over 15,000 ICD-9-CM codes into 804 diagnostic groups, or DxGroups. Each ICD-9-CM code maps to exactly one DxGroup, which represents a well-specified medical condition, such as DxGroup 28.01 Acute Liver Disease. DxGroups are further aggregated into 189 Condition Categories, or CCs.[4] CCs describe a broader set of similar diseases, generally organized into body systems, somewhat like ICD-9-CM major diagnostic categories.

[4] Most CCs are assigned entirely with ICD-9-CM codes. But CCs 185-189 are assigned by beneficiary utilization of selected types of DME, such as wheelchairs. CC 173, Major Organ Transplant, is defined by procedure codes only. CC 129, ESRD is defined by Medicare entitlement status. None of these CCs are included in the CMS-HCC model.

Figure 1

Hierarchical Condition Categories Aggregations of ICD-9-CM Codes



NOTE: ICD-9-CM is *International Classification of Diseases, Ninth Revision, Clinical Modification.*
SOURCE: (Pope et al., 2000b.)

Although they are not as homogeneous as DxGroups, CCs are both clinically- and cost-similar. An example is CC 28 Acute Liver Failure/Disease that includes DxGroups 28.01 and 28.02 Viral Hepatitis, Acute or Unspecified, with Hepatic Coma.

Hierarchies are imposed among related CCs, so that a person is only coded for the most severe manifestation among related diseases. For example (Figure 2), ICD-9-CM Ischemic Heart Disease codes are organized in the Coronary Artery Disease hierarchy, consisting of 4 CCs arranged in descending order of clinical severity and cost, from CC 81 Acute Myocardial Infarction to CC 84 Coronary Athlerosclerosis/Other Chronic Ischemic Heart Disease. A person with an ICD-9-CM code in CC 81 is excluded from being coded in CCs 82, 83, or 84 even if codes that group into those categories were also present. Similarly, a person with ICD-9-CM codes that group into both CC 82 Unstable Angina and Other Acute Ischemic Heart Disease, and CC 83 Angina Pectoris/Old Myocardial Infarction is coded for CC 82, but not CC 83. After imposing hierarchies, CCs become Hierarchical Condition Categories, or HCCs.[5]

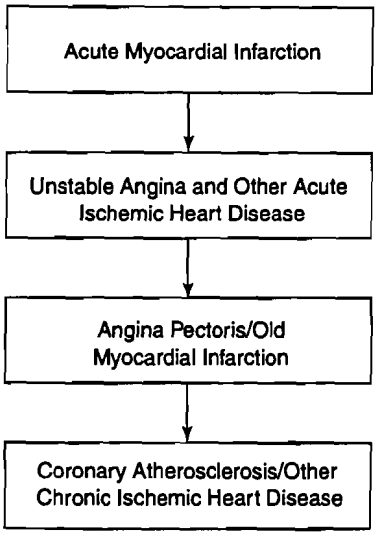
Although HCCs reflect hierarchies among related disease categories, for unrelated diseases, HCCs accumulate. For example, a male with heart disease, stroke, and cancer has (at least) three separate HCCs coded, and his predicted cost will reflect increments for all three problems. The HCC model is more than simply additive because some disease combinations interact. For example, the presence of both Diabetes and Congestive Heart Failure (CHF) could increase predicted cost by more (or less) than the sum of the separate increments for people who have diabetes or CHF alone.

We tested 35 two- and three-way interactions among six common and high-cost chronic diseases defined by HCCs or

[5]The full list of hierarchies used in the CMS-HCC model is available on request from the authors.

Figure 2

Hierarchical Condition Categories Coronary Artery Disease Hierarchy



SOURCE: (Pope et al., 2000b.)

groups of HCCs: diabetes, cerebrovascular disease, vascular disease, or chronic obstructive pulmonary disease (COPD), CHF, and coronary artery disease (Pope et al., 2000b), as well as three interactions of several of these conditions with renal failure.[6] Simple additivity yields most of the explanatory power, in the sense that adding all 38 interactions barely increased the base DCG/HCC model's R^2 (from 11.10 to 11.13 percent). However, six interactions were substantial in magnitude, statistically significant, and clinically plausible. Hence, to improve clinical face validity and predictive accuracy for important subgroups of beneficiaries, we include them in the DCG/HCC model. For example, the simultaneous presence of CHF and COPD leads to higher expected costs than would be calculated by adding the separate increments for CHF and COPD alone.

Because a single beneficiary may be coded for none, one, or more than one DxGroup or HCC, the DCG/HCC model can individually price tens of thousands of distinct clinical profiles using fewer than 200 parameters. The model's structure thus provides, and predicts from, a detailed comprehensive clinical profile for each individual.

HCCs are assigned using hospital and physician diagnoses from any of five sources: (1) principal hospital inpatient; (2) secondary hospital inpatient; (3) hospital outpatient; (4) physician; and (5) clinically-trained non-physician (e.g., psychologist, podiatrist). The DCG/HCC model does not distinguish among sources; in particular, it places no premium on diagnoses from inpatient care. Using Medicare 5-percent sample FFS data, we investigated adding diagnoses from other sources (Pope et al., 2000b). Adding diagnoses from home health providers raised the explanatory power of the base model from

[6] In later work unpublished work, we also examined all two-way interactions of cancer with the other six diagnoses, but did not find any significant effects.

11.15 to 11.65 percent. Further adding diagnoses from DME suppliers raised the explanatory power from 11.65 to 11.85 percent. All other sources of diagnoses either add no predictive power (SNF, ASC, or hospice) or detract from predictive power (clinical laboratory and radiology/imaging clinics). Diagnoses assigned by home health and DME providers are likely to be less reliable than those assigned by physicians or other providers with greater clinical training. Diagnoses from laboratory and imaging tests are also problematic given the significant proportion of rule-out diagnoses. In implementing the CMS-HCC model, potential gains in predictive power from using additional sources were balanced against the costs of collecting and auditing these data; the decision was to only ask MCOs to collect diagnoses from the five baseline sources previously listed.

Consistent with principle 10, we excluded discretionary diagnostic categories (HCCs) from the preliminary prospective payment model. We excluded diagnoses that were vague/non-specific (e.g., symptoms), discretionary in medical treatment or coding (e.g., osteoarthritis), not medically significant (e.g., muscle strain), or transitory or definitively treated (e.g., appendicitis). We also excluded HCCs that did not (empirically) add to costs, and finally, the five HCCs that were defined by the presence of procedures or use of DME, because, as much as possible, we wanted payments to follow what medical problems were present as opposed to what services were offered.[7] Altogether, we excluded 88 of the 189 HCCs, leaving 101 HCCs in the preliminary prospective payment model. As discussed further, additional HCCs were excluded from the final, 70-category CMS-HCC model.

The DCG/HCC model also relies on demographics. Demographic adjusters included in the model are 24 mutually exclusive age/sex cells (e.g., female, age 65-69), an indicator for at least 1-month of Medicaid enrollment in the base year (a poverty indicator), and an indicator of originally disabled status. The age cells distinguish beneficiaries currently entitled to Medicare by age (65 or over) versus disability (under 65); a separate, explicit aged versus disabled entitlement status indicator would be redundant. The originally disabled indicator distinguishes beneficiaries who are currently age 65 or over, but were first entitled to Medicare before age 65 by disability. The age/sex, Medicaid, and originally disabled categories add to each other and to the HCC diagnostic categories.[8] The demographic variables are the same as have been used in the PIP-DCG model, and are discussed at greater length elsewhere (Pope et al., 2000a).

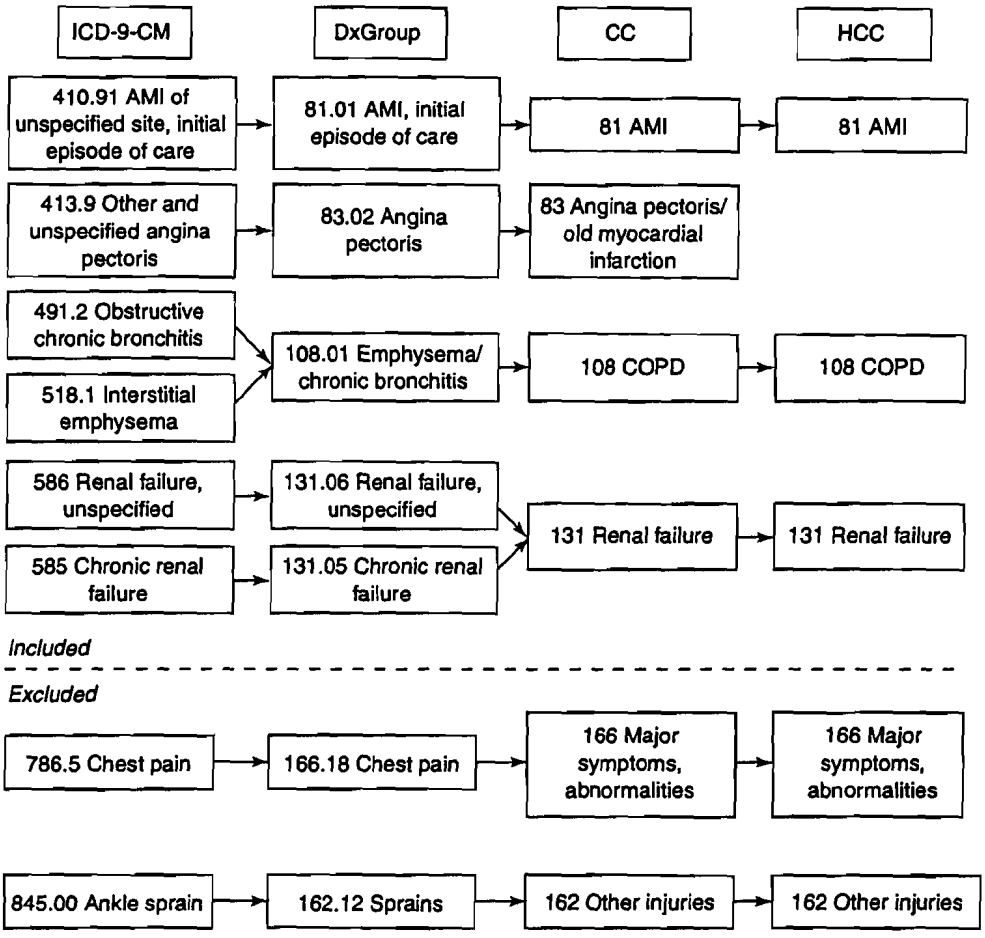
Figure 3 displays a hypothetical clinical vignette of a female age 79, eligible for Medicaid and diagnosed with acute myocardial infarction (AMI), angina pectoris, COPD, renal failure, chest pain, and an ankle sprain. Note that although this female receives CCs for both AMI and angina, she receives no HCC for angina because AMI is a more severe manifestation of coronary artery disease. Also note that while payment includes additive increments for females age 75-79 (demographic categories not shown in Figure 3), Medicaid, AMI, COPD, and renal failure, the HCCs for major symptoms and other injuries are excluded from the payment calculation. Chest pain is a symptom associated with a variety of medical conditions ranging from minor to serious, and sprains are transitory, with minimal implications for next year's cost.

[7] The DME HCCs were developed to predict costs associated with functional impairment not captured by diagnoses. Although they did improve prediction for the functionally impaired, substantial under-prediction remained (Pope et al., 2000b; Kautter and Pope, 2001).

[8] We did not systematically investigate interactions of age and sex with HCCs (diagnoses). This is a subject for future research.

Figure 3

Clinical Vignette for Hierarchical Condition Categories Classification 79 Year Old Female with AMI, Angina Pectoris, COPD, and Renal Failure



NOTES: AMI is acute myocardial infarction. COPD is chronic obstructive pulmonary disease.
SOURCE: (Pope et al., 2000b.)

PERFORMANCE OF DCG/HCC AND PIP-DCG MODELS

The predictive accuracy of risk-adjustment models is typically judged by the R^2 statistic (percentage of variation explained) to measure predictive accuracy for individuals and predictive ratios (ratios of mean predicted to mean actual expenditures for subgroups of beneficiaries) to measure predictive accuracy for groups. The R^2 of age/sex, PIP-DCG, and DCG/HCC models as measured on 1996-1997 Medicare's 5-percent sample FFS data are: age/sex, 1.0 percent; PIP-DCG, 6.2 percent; and DCG/HCC, 11.2 percent.

Adding PIP-DCG to demographic predictors (age/sex) increases predictive power sixfold. Adding secondary inpatient and ambulatory diagnoses (hospital outpatient and physician), and arraying them in a multi-condition cumulative model (DCG/HCC) nearly doubles the power again. Besides the R^2, another interesting summary statistic is the percentage of payments based on demographic variables: 100

Table 1

Predictive Ratios[1] for Alternative Risk-Adjustment Models

Category	Model		
Quintiles of Expenditures	Age/Sex	PIP-DCG	DCG/HCC
First (Lowest)	2.66	2.09	1.23
Second	1.93	1.54	1.23
Third	1.37	1.10	1.14
Fourth	0.95	0.84	1.02
Fifth (Highest)	0.44	0.75	0.86
Top 5 Percent	0.28	0.61	0.77
Top 1 Percent	0.17	0.47	0.69
Hospitalizations			
None	1.33	1.07	1.03
1	0.63	1.02	1.02
2	0.44	0.91	0.98
3 or More	0.26	0.69	0.82
Diagnoses[2]			
Heart Failure	0.47	0.74	0.97
Heart Attack	0.45	0.78	0.98
COPD	0.59	0.79	0.99
Hip Fracture	0.56	0.83	0.99
Depression	0.54	0.77	0.92
Colorectal Cancer	0.60	0.78	0.98
Cerebral Hemorrhage	0.44	0.73	1.04

[1] Mean predicted cost divided by mean actual cost.

[2] From either inpatient or ambulatory setting.

NOTES: Expenditures, hospitalizations, and diagnoses are measured in the base year. COPD is chronic obstructive pulmonary disease.

SOURCE: (Pope et al., 2000b.)

percent in a demographic model, 81 percent in the PIP-DCG model, but only 43 percent in the DCG/HCC model (Pope et al., 2001). With over one-half of payments determined by diagnoses, the DCG/HCC model moves decisively away from the AAPCC demographic-based payment system.

Table 1 shows predictive ratios for selected groups of Medicare beneficiaries. Ratios close to 1.0 indicate accurate prediction of costs; less than 1.0, under prediction; and, more than 1.0, over prediction. The PIP-DCG model improves substantially on age/sex, and in almost all cases, the DCG/HCC model improves significantly on the PIP-DCG model. This is true even for hospitalizations, where the PIP-DCG model distinguishes between those hospitalized or not, while the DCG/HCC model makes no distinction by source of diagnosis.[9] Despite the DCG/HCC model's impressive gains over the age/sex and PIP-DCG models, it still under-predicts for the most expensive and most often hospitalized beneficiaries.

CMS-HCC MODEL

This section describes how the DCG/HCC model was modified before implementation as the M+C risk adjuster for capitation payments in 2004. We will refer to the modified model as CMS-HCC.

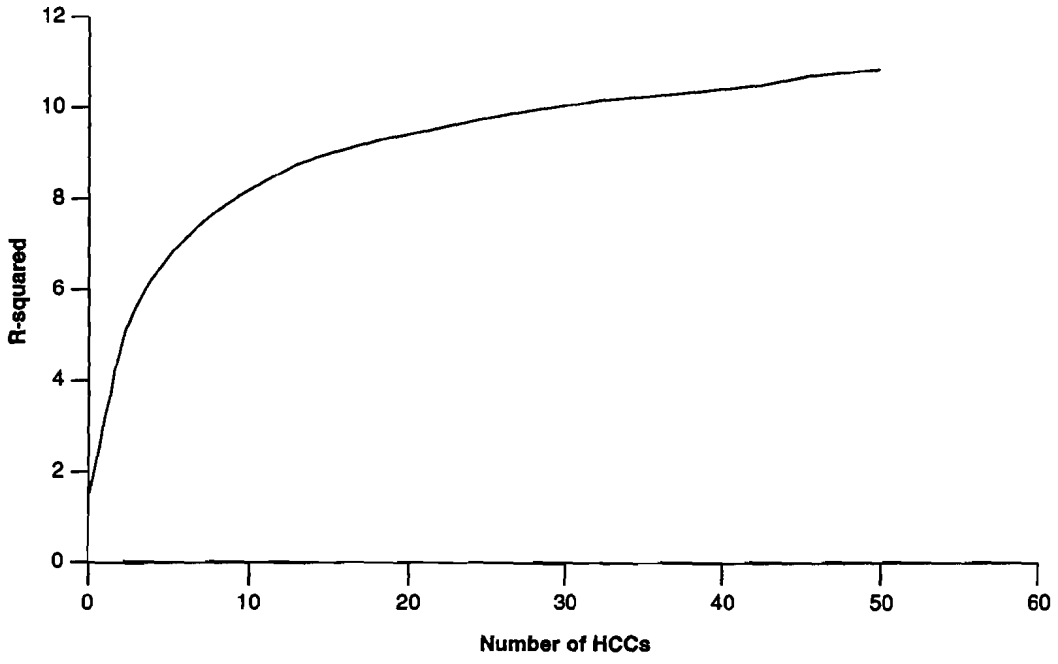
DCG/HCC Model Modification to Simplify Data Collection

When several MCOs withdrew from the M+C program around the year 2000, CMS sought to improve plan retention. Since some MCOs had complained of the burden of collecting encounter data for risk-adjustment, CMS sought to develop risk adjustment models that predict well and rely on ambulatory data, but with reduced data col-

[9] The DCG/HCC model captures multiple conditions that might be diagnosed in multiple inpatient stays, whereas the PIP-DCG model captures only the single principal inpatient diagnosis most predictive of future costs if multiple inpatient stays occur.

Figure 4

Model Explanatory Power as a Function of Number of Hierarchical Condition Categories (HCC)



NOTES: All models, including the one with zero HCCs, include 24 age/sex cells, and Medicaid and originally disabled status. Results based on stepwise regression analysis.

SOURCE: (Pope et al., 2001.)

lection requirements. One measure of the data collection burden imposed by a model is its number of diagnostic categories.[10]

We investigated the relationship between number of diagnostic categories used in the DCG/HCC model and its predictive power (Pope et al., 2001). Figure 4 plots the relationship between number of diagnostic categories and model explanatory power measured by R^2. Diagnostic categories (HCCs) were entered into the model in descending order of their incremental explanatory power using stepwise regression. The base model (with zero HCCs) includes 26 demographic variables, the 24 age/sex cells, and Medicaid and originally disabled status. Its R^2 is 1.69 percent.

The incremental contribution to predictive power declines rapidly with the number of diagnostic categories added to the model. The first diagnostic category entered by the stepwise regression is CHF, which more than doubles the demographic model R^2 to 4.11 percent. The second condition category entered is COPD, raising the R^2 to 4.94 percent. This is an incremental gain of 0.83 percentage points, substantial, but much less then the increment of 2.42 percentage points due to CHF. With 5 HCCs included, 61 percent of the maximum explanatory power of the full (101 HCC) model is attained; with 10 HCCs, 74 percent of the maximum is achieved; with 20, 85 percent, and with 30, 90 percent. The incremental R^2 from adding a diagnostic category is 0.48 percentage points at 5 HCCs; 0.26 percentage points at 10 HCCs; 0.08 percentage points at 20 HCCs; and 0.05 percentage points at 30 HCCs.

[10] The relationship between number of diagnostic categories and data collection burden is controversial. Some MCOs seemed to feel that it would be less burdensome to report all diagnoses, which CMS allows.

This analysis shows that a parsimonious risk-adjustment model with a substantially reduced number of diagnostic categories is almost as predictive as a full model. But parsimony has a cost. In limiting the number of conditions that affect payment, many serious, high-cost diagnoses—especially rare ones—will be ignored. MCOs enrolling beneficiaries with excluded diagnoses will be disadvantaged, and beneficiaries with such conditions may not be well served by MCOs.

CMS considered these results, and consulted with clinicians, on the tradeoff between number of diagnostic categories and predictive power, and also other criteria for diagnostic categories to include in risk adjustment, such as well-defined diagnostic criteria and clinical coherence and homogeneity. It was important that the HCC hierarchies not be disrupted by deletion of higher-ranked HCCs while lower-ranked HCCs were retained. After this process, CMS selected 70 HCCs to include in the CMS-HCC model. The choices reflect a balance among the competing considerations of reducing data collection burden, maximizing predictive power, including rare, high-cost conditions, and selecting only well-defined and clinically coherent conditions. Generally, the higher-cost, more severe conditions at the top of the HCC disease hierarchies were retained, while some lower-cost, more frequent and more discretionary conditions at the bottom of the hierarchies were pruned. For example, in the coronary artery disease hierarchy, AMI (heart attack), other acute IHD (e.g., unstable angina), and angina pectoris/old myocardial infarction were retained, but chronic IHD (e.g., coronary atherosclerosis) was excluded.

After the CMS-HCC model was finalized, a list of approximately 3,000 of the more than 15,000 ICD-9-CM diagnosis codes was identified that are sufficient to define the model's 70 HCCs. In addition, because the CMS-HCC model does not give extra credit for multiple reports of the same diagnosis, MCOs need only report a single encounter during the relevant year of data collection that establishes the diagnosis. The information required for the single encounter is: (1) beneficiary identification number, (2) date (to establish that the diagnosis was made during the relevant reporting period), (3) setting (to establish that the diagnosis was made in one of the allowed hospital or physician settings), and (4) ICD-9-CM diagnosis code. In short, MCOs are required to report only the minimum.

Concern about the quality of diagnostic reporting is the greatest in physician offices, where diagnoses have not heretofore affected payment, and recording of diagnoses is less rigorously practiced than in hospitals. The auditing standard that CMS has promulgated for reporting of physician office diagnoses is that a physician has established the diagnosis in the medical record, and that medical coders have recorded it in accordance with ICD-9-CM rules. CMS will conduct coding audits, but not clinical audits. That is, CMS will require MCOs to demonstrate that a diagnosis is present in the medical record on the specified date and has been coded according to ICD-9-CM. CMS will not require clinical verification of these diagnoses, such as diagnostic test results.

CMS-HCC Model Calibration

We calibrated the CMS-HCC model to 1999-2000 Medicare 5-percent sample FFS data for beneficiaries entitled by age or disability (beneficiaries entitled by ESRD were excluded). The model is prospective, meaning that diagnoses collected in a base year (1999) are used to predict expenditures in the following year (2000). An

important operational change from the PIP-DCG model is that the data lag will be eliminated, making the application of the model consistent with its calibration. With the PIP-DCG model, the data collection period for a payment year ended 6 months before the start of the year, i.e., on June 30 of the previous year, so that final capitation rates could be published by January 1 of the payment year. With the CMS-HCC model, provisional rates will be established by January 1 based on 6-month lagged data, and final rates will be available by June 30 of the payment year based on the previous calendar year's diagnoses. A reconciliation process will adjust the first 6 months of payments to the final rates, if necessary.

A standard set of sample restrictions was employed to ensure a population of beneficiaries with complete 12-month base year diagnostic profiles and complete payment year Medicare expenditures from the FFS claims for aged and disabled beneficiaries (Pope et al., 2000b). Decedents are included in the payment year for their eligible period. Complete FFS claims are not available for months of M+C enrollment or when Medicare is a secondary payer, and M+C plans are not responsible for hospice care, so these months were excluded from our sample. The final sample size is 1,337,887 beneficiaries.

We summed all Medicare payments for a beneficiary for months in 2000 satisfying our sample restrictions, excluding (1) deductibles and copayments paid by the beneficiary; (2) hospice payments; and (3) indirect medical education payments. Hospice and indirect medical education payments are excluded because they were not included in M+C capitation rates, but were paid directly to hospices and teaching hospitals utilized by M+C enrollees. Payments were annualized by dividing them by the fraction of months in 2000 that satisfy our sample restrictions; all analyses are weighted by this eligibility fraction. In general, annualization and weighting ensures that monthly payments are correctly estimated for all beneficiaries, including those who died (Ellis et al., 1996).[11]

The model was calibrated using weighted least squares multiple regression. The CMS-HCC regression model estimated for the combined aged and disabled Medicare population is shown in Table 2.

The elements of the model are:

- Age/sex cells (24).
- Medicaid interacted with sex and age/disabled entitlement status.
- Originally disabled status interacted with sex.
- HCC diagnostic categories (70).
- Interactions of diagnostic categories with entitlement by disability (5).
- Disease interactions (6).

The R^2 for this model is 9.8 percent. Several coefficients are constrained because the unconstrained coefficients violate the principle that higher-ranked conditions in a hierarchy should have higher predicted costs, or for other reasons.[12]

As an example of expenditure prediction, consider our hypothetical scenario in Figure 3 of a female age 79 eligible for Medicaid diagnosed with AMI, angina pectoris, COPD, renal failure, chest pain, and an ankle sprain. The female receives the following incremental cost predictions: female, 75 to 79, $2,562; aged, female, Medicaid, $616; AMI (HCC 81), $1,885; angina pectoris, $0; COPD (HCC 108), $1,936; renal failure (HCC 131), $2,908;

[11] In our calibration, we did not make any geographic adjustments to Medicare payments. In past work, we have found that deflating payments by a geographic input price index had little effect on estimated risk-adjustment model parameters.

[12] Clinical consultants to CMS suggested that metastatic cancer is not consistently correctly coded, so HCCs 7 and 8 were constrained to have equal coefficients. HCCs 81 and 82 were constrained to have equal coefficients because the ICD-9-CM diagnostic detail CMS collects from health plans is not sufficient to distinguish them.

Table 2

Centers for Medicare & Medicaid Services-Hierarchical Condition Categories (CMS-HCC) Combined, Community, and Institutional Models

		Models					
		Combined		Community		Institutional	
Number of Observations		1,337,887		1,291,308		65,593	
R^2		0.0977		0.0976		0.0596	
Adjusted R^2		0.0977		0.0976		0.0589	
Dependent Variable Mean		5,352		5,213		8,937	
Root Mean Square Error		13,407		13,337		15,954	
Model Parameters		105		105		50	
Variable		Parameter Estimate	*t*-ratio	Parameter Estimate	*t*-ratio	Parameter Estimate	*t*-ratio
Female							
0-34 Years		678	3.81	598	3.36	5,457	11.72
35-44 Years		1,110	8.82	1,012	8.03	5,457	11.72
45-54 Years		1,177	11.20	1,096	10.40	5,457	11.72
55-59 Years		1,463	11.87	1,360	11.00	5,457	11.72
60-64 Years		1,996	17.26	1,924	16.56	5,457	11.72
65-69 Years		1,648	42.11	1,572	40.15	5,970	11.73
70-74 Years		2,061	60.25	1,970	57.42	6,049	17.09
75-79 Years		2,562	71.59	2,475	68.56	5,089	19.63
80-84 Years		2,998	71.39	2,936	68.34	4,813	22.51
85-89 Years		3,360	63.45	3,408	61.01	4,515	23.28
90-94 Years		3,683	46.81	4,077	46.25	4,048	19.08
95 Years or Over		3,128	23.27	4,130	25.32	2,980	10.34
Male							
0-34 Years		405	2.72	346	2.32	5,664	13.77
35-44 Years		701	6.63	617	5.81	5,664	13.77
45-54 Years		1,059	12.15	973	11.14	5,664	13.77
55-59 Years		1,460	13.42	1,386	12.68	5,664	13.77
60-64 Years		1,824	17.90	1,755	17.13	5,664	13.77
65-69 Years		1,827	41.47	1,774	40.28	7,435	13.24
70-74 Years		2,380	59.66	2,323	58.17	6,350	14.34
75-79 Years		3,031	69.04	2,960	67.13	6,210	16.45
80-84 Years		3,454	62.03	3,372	59.83	6,201	17.67
85-89 Years		4,129	52.24	4,050	49.80	6,366	17.40
90-94 Years		4,505	32.20	4,620	31.08	5,378	11.29
95 Years or Over		4,753	15.83	5,307	15.89	4,287	5.34
Medicaid and Originally Disabled Interactions with Age and Sex							
Medicaid-Female-Disabled		1,141	11.31	1,133	11.18	—	—
Medicaid-Female-Aged		616	12.91	940	18.18	—	—
Medicaid-Male-Disabled		632	6.80	592	6.31	—	—
Medicaid-Male-Aged		788	10.33	944	11.62	—	—
Originally Disabled-Female		1,231	17.34	1,213	16.44	—	—
Originally Disabled-Male		809	11.66	757	10.73	—	—
Disease Coefficients	**Label**						
HCC1	HIV/AIDS	3,587	13.16	3,514	12.88	6,893	5.42
HCC2	Septicemia/Shock	4,365	34.74	4,563	32.92	4,854	13.89
HCC5	Opportunistic Infections	3,643	10.43	3,346	9.29	6,893	5.42
HCC7	Metastatic Cancer and Acute Leukemia	7,438	81.16	7,510	81.00	2,771	4.54
HCC8	Lung, Upper Digestive Tract, and Other Severe Cancers	7,438	81.16	7,510	81.00	2,771	4.54
HCC9	Lymphatic, Head and Neck, Brain, and Other Major Cancers	3,540	35.91	3,539	35.51	2,319	3.50
HCC10	Breast, Prostate, Colorectal and Other Cancers and Tumors	1,209	26.35	1,194	25.79	1,330	4.01

C1

C1

Refer to NOTES at end of table.

Table 2—Continued

Centers for Medicare & Medicaid Services-Hierarchical Condition Categories (CMS-HCC) Combined, Community, and Institutional Models

Variable		Models: Combined Parameter Estimate	Combined t-ratio		Community Parameter Estimate	Community t-ratio		Institutional Parameter Estimate	Institutional t-ratio	
Disease Coefficients	**Label**									
HCC15	Diabetes with Renal or Peripheral Circulatory Manifestation	3,827	37.71		3,921	36.90		3,137	10.49	
HCC16	Diabetes with Neurologic or Other Specified Manifestation	2,931	30.09		2,833	28.43		3,137	10.49	
HCC17	Diabetes with Acute Complications	2,056	7.84		2,008	7.41		3,137	10.49	
HCC18	Diabetes with Ophthalmologic or Unspecified Manifestation	1,839	18.35		1,760	17.32		3,137	10.49	
HCC19	Diabetes without Complication	1,055	26.10		1,024	25.02		1,308	5.32	
HCC21	Protein-Calorie Malnutrition	3,818	27.52		4,727	29.77		2,193	6.49	
HCC25	End-Stage Liver Disease	4,496	14.91		4,616	14.92		1,375	5.09	
HCC26	Cirrhosis of Liver	2,727	11.93		2,645	11.37		1,375	5.09	
HCC27	Chronic Hepatitis	1,839	6.73		1,841	6.71		1,375	5.09	
HCC31	Intestinal Obstruction/ Perforation	1,997	21.69		2,094	21.62		1,375	5.09	
HCC32	Pancreatic Disease	2,336	17.30		2,281	16.61		1,375	5.09	
HCC33	Inflammatory Bowel Disease	1,574	10.25		1,575	10.16		1,375	5.09	
HCC37	Bone/Joint/Muscle Infections/ Necrosis	2,629	19.68		2,546	18.41		2,539	4.42	
HCC38	Rheumatoid Arthritis and Inflammatory Connective Tissue Disease	1,683	27.72		1,653	26.93		1,463	3.61	
HCC44	Severe Hematological Disorders	5,055	30.80		5,188	30.69		2,299	4.08	
HCC45	Disorders of Immunity	4,224	26.77		4,260	26.64		2,299	4.08	
HCC51	Drug/Alcohol Psychosis	1,571	6.57		1,810	6.99		1,131	6.06	
HCC52	Drug/Alcohol Dependence	1,477	6.15		1,361	5.44		1,131	6.06	
HCC54	Schizophrenia	2,592	26.75		2,786	27.04		1,131	6.06	
HCC55	Major Depressive, Bipolar, and Paranoid Disorders	2,024	30.00		2,209	30.85		1,131	6.06	
HCC67	Quadriplegia, Other Extensive Paralysis	5,665	27.45		6,059	27.20		504	3.94	
HCC68	Paraplegia	5,665	27.45		6,059	27.20		504	3.94	
HCC69	Spinal Cord Disorders/ Injuries	2,484	17.77		2,526	17.45		504	3.94	
HCC70	Muscular Dystrophy	2,239	3.82		1,981	3.27		504	3.94	
HCC71	Polyneuropathy	1,480	19.74		1,377	18.06		504	3.94	
HCC72	Multiple Sclerosis	2,329	11.44		2,654	12.19		504	3.94	
HCC73	Parkinson's and Huntington's Diseases	1,954	19.69		2,436	22.04		504	3.94	
HCC74	Seizure Disorders and Convulsions	1,334	17.25		1,381	16.68		504	3.94	
HCC75	Coma, Brain Compression/ Anoxic Damage	2,396	7.88	C1	2,912	8.62	C1	504	3.94	C2
HCC77	Respirator Dependence/ Tracheostomy Status	10,417	29.54		10,783	28.46		7,259	8.19	
HCC78	Respiratory Arrest	7,543	20.23		7,327	18.79		7,259	8.19	
HCC79	Cardio-Respiratory Failure and Shock	3,451	42.70		3,550	42.39		1,481	4.31	
HCC80	Congestive Heart Failure	2,055	38.48		2,141	38.54		903	4.16	
HCC81	Acute Myocardial Infarction	1,885	31.23		1,785	29.13		1,476	5.75	
HCC82	Unstable Angina and Other Acute Ischemic Heart Disease	1,885	31.23		1,785	29.13		1,476	5.75	
HCC83	Angina Pectoris/Old Myocardial Infarction	1,246	22.82		1,205	21.76		1,476	5.75	

Refer to NOTES at end of table.

Table 2—Continued

Centers for Medicare & Medicaid Services-Hierarchical Condition Categories (CMS-HCC) Combined, Community, and Institutional Models

Variable		Models						
		Combined		Community		Institutional		
		Parameter Estimate	t-ratio	Parameter Estimate	t-ratio	Parameter Estimate	t-ratio	
Disease Coefficients	**Label**							
HCC92	Specified Heart Arrhythmias	1,362	31.73	1,363	30.95	961	4.62	
HCC95	Cerebral Hemorrhage	1,901	10.05	2,011	9.88	774	4.01	
HCC96	Ischemic or Unspecified Stroke	1,498	20.90	1,569	20.34	774	4.01	
HCC100	Hemiplegia/Hemiparesis	1,678	13.96	2,241	16.61	504	3.94	
HCC101	Cerebral Palsy and Other Paralytic Syndromes	767	3.34	840	3.42	504	3.94	C2
HCC104	Vascular Disease with Complications	3,432	36.22	3,473	35.49	2,612	6.30	
HCC105	Vascular Disease	1,662	39.94	1,832	41.72	583	3.72	
HCC107	Cystic Fibrosis	1,936	45.73	1,929	44.87	1,180	4.69	
HCC108	Chronic Obstructive Pulmonary Disease	1,936	45.73	1,929	44.87	1,180	4.69	
HCC111	Aspiration and Specified Bacterial Pneumonias	3,010	20.47	3,556	21.53	2,377	6.82	
HCC112	Pneumococcal Pneumonia, Empyema, Lung Abscess	1,151	6.55	1,034	5.68	2,377	6.82	
HCC119	Proliferative Diabetic Retinopathy and Vitreous Hemorrhage	1,975	13.36	1,791	11.96	5,102	5.46	
HCC130	Dialysis Status	15,926	26.97	15,778	25.96	15,959	5.82	
HCC131	Renal Failure	2,908	23.20	2,954	22.73	2,152	6.26	
HCC132	Nephritis	1,541	6.95	1,401	6.23	2,152	6.26	
HCC148	Decubitus Ulcer of Skin	3,888	32.32	5,285	37.28	1,628	5.98	
HCC149	Chronic Ulcer of Skin, Except Decubitus	2,381	26.76	2,485	26.65	1,346	3.98	
HCC150	Extensive Third-Degree Burns	4,427	2.36	4,935	2.54	1,274	3.37	
HCC154	Severe Head Injury	2,396	7.88	C1 2,912	8.62	C1 1,274	3.37	
HCC155	Major Head Injury	1,211	8.43	1,239	8.08	1,274	3.37	C3
HCC157	Vertebral Fractures w/o Spinal Cord Injury	2,462	20.64	2,514	20.23	504	3.94	C2
HCC158	Hip Fracture/Dislocation	1,301	13.37	2,010	18.51	0	—	
HCC161	Traumatic Amputation	3,965	17.86	C2 4,322	17.92	C2 1,274	3.37	C3
HCC164	Major Complications of Medical Care and Trauma	1,438	18.25	1,346	16.60	1,347	3.66	
HCC174	Major Organ Transplant Status	3,790	8.55	3,702	8.37	4,523	11.13	
HCC176	Artificial Openings for Feeding or Elimination	3,810	23.84	4,054	22.39	4,523	11.13	
HCC177	Amputation Status, Lower Limb/Amputation Complications	3,965	17.86	C2 4,322	17.92	C2 1,274	3.37	C3
Disabled/Disease Interactions								
D-HCC5	Disabled Opportunistic Infections	3,965	5.49	4,047	5.52	—	—	
D-HCC44	Disabled Severe Hematological Disorders	4,649	9.98	4,580	9.72	—	—	
D-HCC51	Disabled Drug/Alcohol Psychosis	2,830	7.12	2,608	6.32	—	—	
D-HCC52	Disabled Drug/Alcohol Dependence	2,160	6.90	2,122	6.61	—	—	
D-HCC107	Disabled Cystic Fibrosis	9,691	6.70	9,547	6.63	—	—	

Refer to NOTES at end of table.

Table 2—Continued

Centers for Medicare & Medicaid Services-Hierarchical Condition Categories (CMS-HCC) Combined, Community, and Institutional Models

Variable		Models					
		Combined		Community		Institutional	
		Parameter Estimate	*t*-ratio	Parameter Estimate	*t*-ratio	Parameter Estimate	*t*-ratio
Disease Interactions							
INT1	DM-CHF1	1,265	14.62	1,296	14.46	1,064	2.91
INT2	DM-CVD	490	4.05	639	4.89	—	—
INT3	CHF-COPD	1,261	14.82	1,238	14.06	1,906	4.95
INT4	COPD-CVD-CAD	316	1.49	406	1.82	—	—
INT5	RF-CHF1	857	3.94	1,202	5.24	—	—
INT6	RF-CHF-DM1	4,185	18.48	4,433	18.71	—	—

NOTES: Beneficiaries with the three-way interaction RF-CHF-DM are excluded from the two-way interactions DM-CHF and RF-CHF. DM is diabetes mellitus (HCCs 15-19). CHF is congestive heart failure (HCC 80). COPD is chronic obstructive pulmonary disease (HCC 108). CVD is cerebrovascular disease (HCCs 95-96, 100-101). CAD is coronary artery disease (HCCs 81-83). RF is renal failure (HCC 131). "|" means coefficients of HCCs are constrained to be equal. C1, C2, and C3 denote non-contiguous constraints.

SOURCE: Pope, G.C. and Kautter, J., RTI International, Ellis, R.P. and Ash, A.S., Boston University, Ayanian, J.Z., Harvard Medical School and Brigham and Women's Hospital, Iezzoni, L.I., Harvard Medical School and Beth Israel Deaconess Medical Center, Ingber, M.J., Levy, J.M., and Robst, J., Centers for Medicare & Medicaid Services, Analysis of 1999-2000 Medicare 5% Standard Analytic File (SAF).

chest pain, $0; and ankle sprain, $0[13] (Table 2). Her total cost prediction is the sum of these increments, or $9,907.

Calibration of DCG/HCC models on several years of data reveals increasingly thorough diagnostic coding. For example, if 1999 diagnoses are used to predict expenditures with a model calibrated on 1996/1997 data, mean expenditures will be over predicted. If more complete coding over time is not accounted for, MCOs will be overpaid by the use of current diagnoses with a model calibrated on historical data. CMS makes a slight downward adjustment in HCC-predicted expenditures to account for this.

CMS-HCC Models for Subpopulations

Medicare beneficiaries differ along characteristics that are important for risk adjustment. First, they may be entitled to Medicare in one of three ways: age, disability, or ESRD. Second, some beneficiaries reside in institutions rather than in the community. Third, some enrollees are new to Medicare and do not have complete diagnostic data. Fourth, Medicare is a secondary payer for some beneficiaries. To account for the different cost and diagnostic patterns of these disparate subgroups of beneficiaries, the CMS-HCC model was adapted for Medicare subpopulations. This section describes models for subpopulations.[14]

Beneficiaries Entitled by Disability

Approximately 12 percent of Medicare beneficiaries are entitled to Medicare because they are under age 65 and have a medical condition that prevents them from working (the disabled). Models calibrated on the full Medicare population (excluding ESRD eligibles), mostly reflect cost patterns among the elderly, the other 88 percent of the population. The implications of some diagnoses might differ between the elderly and disabled. For example, a diagnosis that is disabling may be more severe, and the cost of treating a disease may vary by age. We considered allowing differences in incremental expenditure weights for some diagnoses (HCCs) for the disabled (Pope et al., 1998; 2000b).

[13] The female receives no incremental cost prediction for angina pectoris because AMI is higher-ranked in the coronary artery disease hierarchy and excludes angina. No incremental prediction is made for chest pain and ankle sprain because these diagnoses are not included in the CMS-HCC model.

[14] Risk-adjustment models for ESRD-entitled and functionally-limited beneficiaries are not described in this article.

Using Medicare's 5-percent sample FFS data (1996-1997), we estimated the DCG/HCC model separately on aged and disabled subsamples. We evaluated differences in age versus disabled parameter estimates according to their statistical significance, magnitude, clinical plausibility, and frequency of occurrence in the disabled population (Pope et al., 2000b). Based on these considerations, we chose nine diagnostic categories to receive incremental payments when they occur among disabled beneficiaries. Five of these categories remained significantly different for the disabled when the CMS-HCC model was re-estimated on 1999-2000 data: opportunistic infections, severe hematological disorders (e.g., hemophilia, sickle cell anemia), drug/alcohol psychosis, drug/alcohol dependence, and cystic fibrosis. Incremental annual payments for these conditions among the disabled (in addition to base payments for the elderly) are substantial, ranging from $2,160 to $9,691.

Other than for these five conditions, disease risk-adjustment weights are the same for the aged and disabled populations. The CMS-HCC model is estimated on a combined sample of aged and disabled beneficiaries, with disabled interactions for these five diagnostic categories. The combined aged/disabled model is shown in Table 2.

Community and Institutional Residents

Using the newly available Medicare MDS, we identified long–term nursing home residents in the current (i.e., payment) year. Long-term nursing home residence was defined as continuously residing in a nursing home for at least 90 days, as indicated by a 90-day clinical assessment reported by the nursing facility through the MDS. In our prospective risk-adjustment modeling sample of 1,337,887 beneficiaries, 65,593 beneficiaries, or 5 percent, had at least 1 month of long-term nursing facility residence in 2000.[15]

Table 3 compares sample sizes and mean expenditures by demographic categories for community and institutional residents, and shows predictive ratios from the CMS-HCC model calibrated on the combined community/institutional sample (Table 2). Nearly one-half (49 percent) of long-term nursing facility residents are age 85 or over. Facility residents are only 2 percent of the combined community plus institutional population for females age 70 to 74, but fully 37 percent of the combined population for females age 95 or over.

Overall, institutional residents are 71 percent more expensive than community residents, $8,937 in mean annualized expenditures compared to $5,213. The age profiles of expenditures are quite different. Among community residents, mean expenditures rise steadily with age in the under 65 disabled population and then again in the elderly population, except for a slight decline for the oldest females. In contrast, among the institutionalized, mean expenditures are fairly constant across all ages until they decline significantly among the oldest old. For all age/sex cells except the oldest old, mean expenditures for the institutionalized are substantially higher than for community-dwelling beneficiaries.

However, although not shown in Table 3, among beneficiaries diagnosed with particular HCCs, mean expenditures for the institutionalized are often similar to those of community residents. For example, among all beneficiaries with CHF (HCC 80), expenditures for the institutionalized are $11,719, which is $255 less than for community residents. More generally, when classifying people by the presence of

[15] Beneficiaries with both community and long-term institutional months in the same year are included in both samples, weighted by the fraction of their total months alive in the year in each status.

Table 3

Descriptive Statistics for Community and Institutionalized Residents

Variable	Community			Institutional		
	Observations	Mean Annualized Expenditures	Predictive Ratio[1]	Observations	Mean Annualized Expenditures	Predictive Ratio[1]
Overall	1,291,308	5,213	0.99	65,593	8,937	1.12
Demographics						
Female						
0-34 Years	7,007	3,623	1.00	49	9,251	0.99
35-44 Years	15,566	4,332	1.00	199	9,395	0.94
45-54 Years	22,077	4,692	1.00	473	8,869	1.07
55-59 Years	14,023	5,254	1.00	343	10,168	0.91
60-64 Years	15,793	5,993	1.00	501	9,906	1.04
65-69 Years	129,970	3,714	1.00	1,380	10,961	0.99
70-74 Years	171,775	4,372	1.00	3,098	10,901	0.97
75-79 Years	157,586	5,260	1.00	6,260	9,458	1.08
80-84 Years	111,303	6,101	0.99	9,801	8,797	1.13
85-89 Years	66,301	6,882	0.97	12,294	8,054	1.19
90-94 Years	26,852	7,606	0.92	9,535	7,146	1.29
95 Years or Over	8,074	7,338	0.83	4,729	5,734	1.42
Male						
0-34 Years	10,272	2,868	1.00	106	10,622	0.95
35-44 Years	22,913	3,666	1.00	384	9,596	0.92
45-54 Years	29,377	3,968	1.00	606	10,186	0.91
55-59 Years	16,391	4,651	1.00	438	10,340	0.96
60-64 Years	18,581	5,214	1.00	588	10,486	1.00
65-69 Years	105,856	4,018	1.00	1,132	12,432	0.88
70-74 Years	128,874	5,014	1.00	1,921	11,501	0.99
75-79 Years	106,402	6,207	1.00	2,842	11,411	1.04
80-84 Years	64,263	7,083	1.00	3,404	11,049	1.06
85-89 Years	30,765	8,144	0.99	3,116	10,754	1.08
90-94 Years	9,343	8,731	0.97	1,783	9,489	1.20
95 Years or Over	1,944	9,062	0.92	611	8,096	1.37
Medicaid	196,604	6,523	0.97	33,074	8,895	1.17
Originally-Disabled	81,894	7,614	0.99	7,415	10,606	1.11

[1] Ratio of mean expenditures predicted by the Centers for Medicare & Medicaid Services - Hierarchical Condition Categories (CMS-HCC) model for combined community/institutional samples to mean actual expenditures.

SOURCE: Pope, G.C. and Kautter, J., RTI International, Ellis, R.P. and Ash, A.S., Boston University, Ayanian, J.Z., Harvard Medical School and Brigham and Women's Hospital, Iezzoni, L.I., Harvard Medical School and Beth Israel Deaconess Medical Center, Ingber, M.J., Levy, J.M., and Robst, J., Centers for Medicare & Medicaid Services, Analysis of 1999-2000 Medicare 5% Standard Analytic File (SAF).

a single diagnosis, expenditures for the institutionalized may be higher, lower, or about the same.

Thus, the main reason that people in facilities cost more is that they have more medical problems, a distinction that is fully accounted for by the HCCs. In fact, the predictive ratios from the combined CMS-HCC model for community and institutional beneficiaries are, respectively, 0.99 and 1.12 (Table 3). This means that the combined model, on average, under predicts expenditures for community residents by 1 percent, and over predicts expenditures for long-term nursing home residents by 12 percent. Lower expenditures among facility residents adjusting for disease burden could result from substituting non-Medicare for Medicare-reimbursed services; since most nursing home service are not reimbursed by Medicare. Also, greater monitoring of nursing home than community residents may identify and prevent problems leading to hospitalization. The under-prediction for community residents and over-prediction for facility residents is most severe for the oldest age groups, most likely due to decisions to limit

aggressive care for very old residents in nursing homes. The over-prediction of the costs of the institutionalized, together with their different cost patterns by age and diagnosis, led us to consider differentiating the CMS-HCC model for community and institutional populations.

Within a multiple regression model estimation framework, we investigated alternative approaches to allowing differences in the model between community and institutional residents, ultimately choosing to estimate separate models. This properly calibrates the prediction of each group's costs, while allowing all demographic and disease coefficients to differ between community and institutional populations.

In addition to the combined model, Table 2 shows the CMS–HCC community and institutional models. Not surprisingly, the community model R^2 and most of the demographic and disease coefficients are very similar to the combined model, because community residents comprise 95 percent of the combined sample. A few coefficients show greater differences. The community coefficients for the oldest age cells are significantly larger than the combined model coefficients because the lower-cost very old institutionalized have been removed from these cells. The community coefficients for the aged enrolled in Medicaid are also significantly higher, as are several HCC coefficients.

The institutional model R^2 is considerably lower than the community model. But some of the community model's predictive power comes from distinguishing beneficiaries who are healthy (no diagnoses) versus sick (with diagnoses), while the institutional model is explaining cost variations among a population comprised entirely of impaired individuals. Diagnoses help explain why someone might be institutionalized (i.e., distinguish healthy from sick), but are not as powerful in explaining expenditure differences among the institutionalized. Disease (HCC) coefficients tend to be smaller in the institutional model than in the community model (Table 2). Diagnoses are less predictive of incremental costs among the more uniformly expensive institutional population than they are among the community population.

We constrained certain groups of demographic and diagnostic coefficients in the institutional model to be equal (Table 2), because the small available sample of institutionalized beneficiaries resulted in their low prevalence in some diagnostic categories (HCCs) and made it difficult to obtain stable estimates of each separate parameter. For the same reason, we included no disabled interaction terms, and only two of the disease interaction terms in the institutional model. Also, HCC 158 Hip Fracture/Dislocation was excluded because its coefficient was negative.

The age/sex coefficients for the institutionalized are much higher than for community residents except for the oldest ages. This implies that institutionalized beneficiaries are predicted to be expensive regardless of their diagnostic profile (e.g., even lacking any of the diagnoses included in the CMS-HCC model), whereas community residents are predicted to be expensive only if diagnosed with at least one of the serious diseases included in the CMS-HCC model. This makes sense since institutionalization itself is a marker of poor health, aside from diagnostic profile, but the institutionalized age/sex coefficients decline for the oldest ages, and fall below the community coefficients. Medical treatment may be less aggressive for old, frail beneficiaries who are institutionalized.

Among the institutional population, the coefficient for Medicaid was negative and the coefficients for originally disabled was statistically insignificant. These variables

were excluded from the institutional model. Beneficiaries often qualify for Medicaid after spending down their personal assets to pay for a lengthy nursing home stay. Thus, Medicaid may be a proxy for beneficiaries in the later portion of their stays, when they are less expensive than in the earlier, post-acute phase of their nursing home tenure.

New Medicare Enrollees

The CMS–HCC model requires a complete 12-month base year diagnostic profile to predict the next year's expenditures. Beneficiaries without 12 months base year Medicare enrollment, but at least 1 month of prediction year enrollment, are defined as new enrollees. About two–thirds of new enrollees are age 65.[16] New enrollees may be under age 65 if they become eligible for Medicare by disability; they may be over age 65 if they delay Medicare enrollment or are not originally enrolled in both Parts A and B.[17] We developed a demographic model to predict expenditures for new enrollees who lack the data needed to apply the CMS-HCC model.

Table 4 presents frequencies and mean annualized expenditures from the 5-percent FFS sample data for new enrollees and continuing enrollees. Continuing enrollees are defined as beneficiaries having 12 months of Parts A and B Medicare enrollment in the base year and at least 1 month in the prediction year. For female and male new enrollees age 65, mean annualized expenditures are $2,729 and $2,900, respectively, less than one-half of costs of continuing enrollees ($6,952 for female and $6,055 for male). For almost all new enrollees age 65, the original reason for Medicare entitlement is age.[18] In contrast, continuing enrollees age 65 were originally entitled to Medicare by disability, and hence are much more expensive. For other ages, mean expenditures of new and continuing enrollees are much more similar. To achieve sufficient sample sizes in all age ranges to calibrate the new enrollees model, we merged the new and continuing enrollees samples, which resulted in a sample size of 1,495,225 with mean expenditures of $5,184. For age 65, actual new enrollees dominate the combined sample, and the cost weight reflects their (low) relative costs. Continuing enrollees age 65 are included in the sample to calibrate the originally disabled coefficient for age 65. For other than age 65, the sample is dominated by continuing enrollees, but their costs appear to proxy actual new enrollee costs reasonably well for younger or older ages.

Beneficiaries for Whom Medicare is a Secondary Payer

Working aged beneficiaries are Medicare beneficiaries, age 65 or over, with private group health insurance coverage from their or their spouse's employer. By law, Medicare is a secondary payer for these beneficiaries. The primary private health plan must pay for medical expenses to the extent of its defined benefits. Only if Medicare covers services not covered by the private plan, or has more generous coverage (e.g., lower deductibles or copayments) for Medicare-covered services, is Medicare responsible for payment, and then only to the extent of the difference in

[16] To simplify the new enrollees model, we recoded new enrollees age 64 on February 1 with an original reason for Medicare entitlement of aged to age 65. Thus, the age 65 cell in the new enrollees model combines new enrollees ages 64 and 65 on February 1 of the prediction year whose original reason for entitlement is aged.

[17] For example, a beneficiary might be entitled to Part A (hospital insurance) by age at age 65 or over, but might not pay Part B (physician insurance) premium until an older age.

[18] Some age 65 new enrollees might have originally been entitled to Medicare by disability when under age 65, but then have rejoined the work force and lost their Medicare eligibility, only to re-enroll at age 65.

Table 4

Descriptive Statistics for New and Continuing Medicare Enrollees[1]

Age/Sex Category	New Enrollees[2]		Continuing Enrollees[3]	
	Observations	Mean Annualized Expenditures	Observations	Mean Annualized Expenditures
Female				
0-34 Years	2,540	3,532	7,037	3,653
35-44 Years	3,685	4,341	15,717	4,385
45-54 Years	5,891	4,814	22,431	4,767
55-59 Years	4,029	4,903	14,277	5,354
60-64 Years	3,310	5,705	16,159	6,094
65 Years	58,946	2,729	3,336	6,952
66 Years	1,448	3,319	29,534	3,401
67 Years	845	3,349	31,560	3,684
68 Years	531	3,116	32,578	3,740
69 Years	504	3,608	33,893	3,905
70-74 Years	1,311	4,672	173,829	4,461
75-79 Years	471	5,063	161,843	5,387
80-84 Years	200	6,043	118,144	6,276
85-89 Years	95	8,111	75,186	7,035
90-94 Years	46	5,931	34,135	7,500
95 Years or Over	15	6,457	11,886	6,795
Male				
0-34 Years	3,434	3,089	10,342	2,934
35-44 Years	4,281	3,690	23,172	3,746
45-54 Years	5,820	4,099	29,814	4,074
55-59 Years	4,120	4,603	16,677	4,772
60-64 Years	4,196	4,775	18,986	5,346
65 Years	46,262	2,900	3,940	6,055
66 Years	1,546	3,205	24,472	3,644
67 Years	872	2,976	25,279	3,933
68 Years	570	3,501	25,915	4,145
69 Years	490	3,638	27,009	4,295
70-74 Years	1,223	5,700	130,148	5,087
75-79 Years	429	6,476	108,214	6,307
80-84 Years	144	5,916	66,505	7,231
85-89 Years	63	8,028	32,848	8,326
90-94 Years	19	13,027	10,601	8,827
95 Years or Over	2	3,221	2,420	8,867

[1] Aged and disabled beneficiaries. Excludes working aged and ESRD beneficiaries.

[2] Enrollees with less than 12 months of base year eligibility.

[3] Enrollees with 12 months of base year eligibility.

SOURCE: Pope, G.C. and Kautter, J., RTI International, Ellis, R.P. and Ash, A.S., Boston University, Ayanian, J.Z., Harvard Medical School and Brigham and Women's Hospital, Iezzoni, L.I., Harvard Medical School and Beth Israel Deaconess Medical Center, Ingber, M.J., Levy, J.M., and Robst, J., Centers for Medicare & Medicaid Services, Analysis of 1999-2000 Medicare 5% Standard Analytic File (SAF).

coverage. Medicare expenditures for working aged beneficiaries are lower for this reason, as well as because working may be a proxy for better health.[19] Estimation of a separate model for the working aged is not feasible with the sample sizes available from the Medicare's 5-percent FFS sample. A simple adjustment to CMS–HCC model predictions is a multiplier that scales cost predictions to be lower for these beneficiaries.

We defined the working aged as beneficiaries otherwise satisfying the requirements of our 1999-2000 aged/disabled prospective modeling sample who had at least 1 month of working aged status in the prediction year (2000). There are 19,057 beneficiaries in our working aged sample, or about 1.4 percent as many individuals as in our aged/disabled sample. The mean annualized expenditures of the working aged are $966, less than one-fifth as much as for the aged/disabled community sample ($5,213). The CMS–HCC community

[19] Throughout this section, we use the terms working and working aged to include both those who are actually working, and the spouses of those who are working.

model over-predicts mean working aged expenditures by a factor of 3.66. Essentially, we define the working aged multiplier as the ratio of mean actual to mean predicted expenditures for the working aged sample, where expenditures are predicted by the CMS-HCC community model. With an adjustment for beneficiaries who have a mixture of working aged and non-working-aged months in the payment year, the working aged multiplier is 0.215.

CONCLUSIONS

CMS' adaptation of the DCG/HCC model makes substantially more accurate predictions of medical costs for M+C enrollees than has previously been possible. Its use is intended to redirect money away from MCOs that cherry-pick the healthy, while providing the MCOs that care for the sickest patients the resources to do so. The ultimate purpose of the CMS-HCC payment model is to promote fair payments to MCOs that reward efficiency and encourage excellent care for the chronically ill. The CMS-HCC model will continue to evolve. Additional diagnoses may be needed to predict drug expenditures incurred under the drug benefit enacted by MMA (2003). The model may need to be recalibrated to reflect new treatment patterns and disease prevalence. Diagnosis-based risk adjustment may need to be coordinated with disease management programs and incentives for quality of care.

The model has evolved over two decades of research,[20] with careful attention to clinical credibility, real-world incentives and feasibility tradeoffs. Continuous feedback between government technical staff and policymakers at CMS on the one hand, and research organization and academic researchers on the other, has shaped the CMS-HCC model. Much of the recent research reported in this article has related to adapting the model for Medicare subpopulations. The use of a single modeling framework—the CMS-HCC model—provides unity and organization to the subgroup models with the unique features specific to certain types of beneficiaries. Comprehensive risk adjustment, based on ambulatory as well as inpatient diagnoses, is just beginning to be implemented. Thus, it is too early to tell whether it will achieve its goals. As risk adjustment continues to be incorporated in Medicare payments to MCOs, it will be important to evaluate its impact on these organizations and the beneficiaries they serve, especially organizations that care for the chronically ill and their enrollees. This will tell us a great deal about the feasibility and consequences of matching health care resources to needs.

[20] The DCG line of risk-adjustment research dates back to the report by Ash et al. (1989), based on research begun in 1984.

ACKNOWLEDGMENT

The authors would like to thank Helen Margulis for her exceptional computer programming assistance.

REFERENCES

Ash, A.S., Porell, F., Gruenberg, L., et al.: Adjusting Medicare Capitation Payments Using Prior Hospitalization. *Health Care Financing Review* 10(4):17-29, Summer 1989.

Brown, R.S., Clement, D.G., Hill, J.W., et al.: Do Health Maintenance Organizations Work for Medicare? *Health Care Financing Review* 15(1):7-23, Fall 1993.

Centers for Disease Control and Prevention: *International Classification of Diseases, Ninth Revision, Clinical Modification (IC+CD-9-CM).* Internet address: http://www.cdc.gov/nchs/about/otheract/icd9/abticd9.ht. (Accessed 2004.)

Ellis, R.P., Pope, G.C., Iezzoni, L.I., et al.: Diagnosis-Based Risk Adjustment for Medicare Capitation Payments. *Health Care Financing Review* 17(3):101-128, Spring 1996.

Ellis, R.P., Pope, G.C., Iezzoni, L.I., et al.: Diagnostic Cost Group (DCG) and Hierarchical Coexisting Conditions (HCC) Models for Medicare Risk Adjustment. Final Report to the Health Care Financing Administration under Contract Number 500-92-0020, Delivery Order Number 6. Health Economics Research, Inc. Waltham, MA. April, 1996.

Ellis, R.P., Ash, A.S.: Refinements to the Diagnostic Cost Group Model. *Inquiry* 32(4):1-12, Winter 1995.

Hughes, J.S., Averill, R.F., Eisenhandler, J., et al.: Clinical Risk Groups (CRGs): A Classification System for Risk-Adjusted Capitation-Based Payment and Health Care Management. *Medical Care* 42(1):81-90, January 2004.

Kapur, K., Tseng, C.W., Rastegar, A., et al.: Medicare Calibration of the Clinically Detailed Risk Information System for Cost. *Health Care Financing Review* 25(1):37-54, Fall 2003.

Kautter, J., and Pope, G.C.: Predictive Accuracy of Diagnostic Cost Group (DCG) Risk Adjustment Models. Final Report to the Centers for Medicare & Medicaid Services under Contract Number 500-95-048. Health Economics Research, Inc. Waltham, MA. August 2001.

Kronick, R., Gilmer, T., Dreyfus, T., et al.: Improving Health-Based Payment for Medicaid Beneficiaries: Chronic Illness and Disability Payment System. *Health Care Financing Review* 21(3):29-64, Spring 2000.

Mello, M.M., Stearns, S.C., Norton, E.C., et al.: Understanding Biased Selection in Medicare HMOs. *Health Services Research* 38(3):961-992, June 2003.

Pope, G.C., Kautter, J., Ash, A.S., et al.: Parsimonious Risk Adjustment Models for Medicare. Final Report to the Centers for Medicare & Medicaid Services under Contract Number 500-95-048. Health Economics Research, Inc. Waltham, MA. December, 2001.

Pope, G.C., Ellis, R.P., Ash, A.S., et al.: Principal Inpatient Diagnostic Cost Group Model for Medicare Risk Adjustment. *Health Care Financing Review* 21(3):93-118, Spring 2000a.

Pope, G.C., Ellis, R.P., Ash, A.S., et al.: Diagnostic Cost Group Hierarchical Condition Category Models for Medicare Risk Adjustment. Final Report to the Health Care Financing Administration under Contract Number 500-95-048. Health Economics Research, Inc. Waltham, MA. December, 2000b.

Pope, G.C., Liu, C.F., Ellis, R.P., et al.: Principal Inpatient Diagnostic Cost Group Models for Medicare Risk Adjustment. Final Report to the Health Care Financing Administration under Contract Number 500-95-048. Health Economics Research, Inc. Waltham, MA. February 1999.

Pope, G.C., Adamache, K.A., Walsh, E.G., and et al.: Evaluating Alternative Risk Adjusters for Medicare. *Health Care Financing Review* 20(2):109-129, Winter 1998.

Pope, G.C., Ellis, R.P., Liu, C.F., et al.: Revised Diagnostic Cost Group (DCG)/Hierarchical Coexisting Conditions (HCC) Models for Medicare Risk Adjustment. Final Report to the Health Care Financing Administration under Contract Number 500-95-048. Health Economics Research, Inc. Waltham, MA. February 1998.

Riley, G., Tudor, C., Chiang, Y., et al.: Health Status of Medicare Enrollees in HMOs and Fee-for-Service in 1994. *Health Care Financing Review* 17(4):65-76, Summer 1996.

Weiner, J.P., Dobson, A., Maxwell, S.L., et al.: Risk-Adjusted Medicare Capitation Rates Using Ambulatory and Inpatient Diagnoses. *Health Care Financing Review* 17(3):77-100, Spring 1996.

Reprint Requests: Gregory C. Pope, RTI International, 411 Waverly Oaks Road, Suite 330, Waltham, MA 02452. E-mail address: gpope@rti.org

Available online at www.sciencedirect.com

SCIENCE DIRECT

Preventive Medicine 39 (2004) 969–975

Preventive Medicine

www.elsevier.com/locate/ypmed

Web-based screening and brief intervention for the spectrum of alcohol problems ☆

Richard Saitz, M.D., M.P.H.,[a,b,d,*] Eric D. Helmuth, M.A.,[c] Susan E. Aromaa, M.S.,[c] Anara Guard, M.S.,[c] Marc Belanger, M.A.,[c] and David L. Rosenbloom, Ph.D.[c,d]

[a] *Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Department of Medicine, Boston University School of Medicine and Boston Medical Center, Boston, MA 02118, USA*

[b] *Department of Epidemiology, Boston University School of Public Health, Boston, MA 02118, USA*

[c] *Join Together, Department of Social and Behavioral Sciences, Boston University School of Public Health, Boston, MA 02118, USA*

[d] *Center to Prevent Alcohol-related Problems Among Young People, Boston University School of Public Health, Boston, MA 02118, USA*

Available online 2 June 2004

Abstract

Context. Many persons who drink excessively remain unidentified and do not receive interventions. Screening and intervention using the World Wide Web could make such services more accessible and therefore more widely used.

Objective. To evaluate the use of a novel alcohol screening and brief intervention Web site.

Design. A Web site was developed, posted, and its use was evaluated. We analyzed a sample of visitors who completed alcohol screening over a 14-month period to describe their alcohol use, and their use of portions of the Web site that provide information and referral resources.

Setting. The Internet.

Patients or other participants. Web site visitors, with a focus on visitors who completed an alcohol-screening questionnaire about their own drinking.

Intervention. Brief intervention via the Web site, consisting mainly of feedback, advice, and a menu of change options and referral information.

Main outcome measures. Self-reported drinking amounts and alcohol screening test scores, and utilization of Web site components.

Results. Visitors completed online alcohol screening questionnaires at a rate of 50,711/year of 115,925 visitors/year. In a 14-month period, 39,842 adults completed the questionnaire about their own drinking habits; 66% were men, 90% reported drinking hazardous amounts (per occasion or typical weekly amounts), 88% reported binge (per occasion) drinking, and 55% reported typically exceeding weekly risky drinking limits. Most (65%) had alcohol screening test results (AUDIT $\geq$ 8) consistent with alcohol abuse or dependence; similar proportions of women and men were hazardous drinkers. One-fifth of visitors visited portions of the Web site that provided additional information about alcohol use and referrals. Visitors with possible alcohol abuse or dependence were more likely than those without these disorders to visit a part of the Web site designed for those seeking additional help (33% vs. 8%, $P < 0.0001$).

Conclusions. A well-publicized, easily accessible, research-based screening and intervention Web site can attract many users, most of whom are drinking excessively, and many of whom avail themselves of referral information after receiving individualized feedback.

Keywords: Screening; Alcoholism; Alcohol use; Internet; Brief intervention

Introduction

Alcohol use disorders are costly and a leading cause of disability and death worldwide [1,2]. Brief screening tools can identify people with alcohol problems, and, once identified, those people can receive brief interventions [3,4]. But many people with alcohol use disorders do not seek care nor are they screened even when they contact the health care system [1,5,6]. Barriers to seeking care include

☆ This work was presented in preliminary form at the meeting of the American Public Health Association in Philadelphia, PA on November 13, 2002.

* Corresponding author. Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Boston Medical Center, 91 East Concord Street #200, Boston, MA 02118-2393. Fax: +1-617-414-4676.

E-mail address: rsaitz@bu.edu (R. Saitz).

0091-7435/$ - see front matter © 2004 The Institute For Cancer Prevention and Elsevier Inc. All rights reserved.
doi:10.1016/j.ypmed.2004.04.011

lack of motivation to change, beliefs about treatment, attitudes, fear of discrimination or job loss if detected, and lack of perceived effective access to care [7–9]. Active screening followed by intervention can improve patient outcomes for problem drinkers, and improve access to specialty treatment for those with dependence [3,10]. Despite these available options, many people remain unidentified and untreated [11,12].

Most screening and brief intervention programs have been designed for health care settings, usually paper and pencil questionnaires or in-person interviews [3]. Web-based measures of alcohol use appear to be at least as reliable as these older methods [13]. Using the Internet for screening has the potential to greatly increase the number of people screened and improve access to brief intervention. More than half of all U.S. adults use the Internet and more than half of these Internet users search for health information online (77 million people); 8% of Internet users (9 million people), 14% of young adult Internet users (18–29 years old), and 7% of middle-aged adult users (30–49 years old) have searched for information on alcohol or drug problems [14].

AlcoholScreening.org provides online self-assessment tools and health-based information to help individuals identify their own risky drinking patterns or current alcohol problems. The site delivers personalized feedback and helps users locate assistance if they are ready to seek help. In this paper, we describe the development of the Web site designed to anonymously screen adults and provide personalized online feedback regarding alcohol use, and the feasibility of its use for alcohol screening on the Web. After creating and posting the Web site, we tested whether this Web site was used widely, whether it reached a target population of hazardous drinkers, and whether these persons would seek further assistance from the site when appropriate.

Methods

Web site design

AlcoholScreening.org is an anonymous, free online self-screening service to assess an individual's alcohol consumption and its consequences. AlcoholScreening.org was based on the health belief model [15]. According to the core concepts of the health belief model, risky drinkers are more likely to reduce their alcohol consumption or otherwise control their at-risk behavior when they believe the threats are real, the benefits of change are valuable, and the barriers to behavioral changes are lessened. With these concepts in mind, AlcoholScreening.org was created as a vehicle to provide the user with personalized feedback speaking directly to the issues of risks, benefits, and action steps. Additional features were created to provide supplemental information for Web site users wishing to take action.

Visitors to this Web site answer 12 questions about their drinking and an additional question for research purposes, and are presented with results that outline the likelihood that their reported drinking patterns indicate risky or harmful alcohol consumption. Visitors may also access an online library of health information about alcohol consumption and alcohol problems, search a national database of substance abuse treatment facilities, or follow links to additional information online. All users are presented with a disclaimer emphasizing that the Web site does not provide a medical diagnosis and cannot substitute for a full evaluation by a health professional.

Visitors complete an online version of the Alcohol Use Disorders Identification Test (AUDIT) [16]. In addition to the 10 AUDIT questions, users are asked two additional questions to further detail the quantity and frequency of their alcohol consumption, and whether their responses reflect their personal alcohol consumption patterns. These questions read:

- "Thinking about a typical *week*, on how many *days* do you have at least one alcoholic drink? (If you don't drink every week, answer for a typical week in which you do)";
- "Thinking about the past year, what is the greatest number of drinks you've had on any one occasion?"; and
- "Optional: This question is for research purposes only: I am completing this test based upon my own alcohol-use experience OR I am just curious about the test and the related feedback, or answered the questions with someone else in mind."

Upon completion of the online screening test, users receive nonjudgmental feedback based on their AUDIT score [17], the current U.S. Department of Agriculture Dietary guidelines for moderate alcohol consumption [18], and U.S. alcohol consumption norms [19]. Feedback is based on the principles of successful brief interventions [20] and distinguishes between drinking amounts that place people at risk for future consequences, and drinking with consequences that have already occurred [21,22]. All users are advised whether their screening results indicate a likelihood of hazardous or harmful alcohol consumption, and all are informed about current guidelines for low-risk drinking. In addition, links are provided to alcohol and health information, and the site provides the option of searching a national database of substance abuse treatment facilities for local, in-person assistance. A traffic light theme is graphically incorporated into the site to help illustrate the screening results.

Users scoring below eight on the AUDIT and whose alcohol consumption falls within the U.S. Dietary guidelines for moderate alcohol consumption are presented with a "green light" results page. They are told that their alcohol consumption appears to fall within healthy limits, but are warned that there are certain circumstances when any amount of alcohol may not be safe, for example, when operating a vehicle or machinery, while pregnant, or if certain medical conditions exist.

Those scoring below eight on the AUDIT, but who exceed the U.S. Dietary guidelines for weekly or per-occasion alcohol consumption see a "yellow light" results page. They are told that while their results do not suggest that a pattern of excess drinking is currently harming their health, the amount they reported consuming on at least one occasion increases their risk for injury or other immediate consequences, or that the amount they reported consuming per week places them at risk for future, mainly chronic, health consequences. They are encouraged to cut down, abstain, or set a safer personal limit for how much alcohol they consume, and, as in the "green light" scenario, are reminded that for some people and in certain situations, no amount of alcohol is safe. In addition, those who exceed weekly alcohol consumption guidelines receive normative information comparing their alcohol consumption to that of the general adult American population and to the adult population of their gender. A sample of this information based on the data provided by a user is: "More than 95% of the general adult American population, and 91% of men consume fewer drinks per week than you reported consuming."

Users scoring eight or above on the AUDIT are presented with a "red light," and are told that likely their current drinking is hazardous or harmful to their health and well being. This group also receives normative feedback comparing their alcohol consumption to that of the general adult American population and to the adult population of their gender. They are told that the AUDIT cannot diagnose any condition or tell them for certain if alcohol use is harming their health, and are advised to seek further evaluation from their doctor or other qualified health professional.

At any time, users can follow links to a national alcohol treatment facility database and information on alcohol and health. The treatment database is a current copy of a data set maintained by the Center for Substance Abuse Treatment (CSAT) of the federal Substance Abuse and Mental Health Services Administration (SAMHSA). Further information on alcohol and health is republished by AlcoholScreening.org and provided by U.S. Department of Health and Human Services (U.S. Department of Agriculture and the SAMHSA Center for Substance Abuse Prevention) and the National Institute on Alcohol Abuse and Alcoholism (NIAAA) of the National Institutes of Health (NIH) [18,23,24].

AlcoholScreening.org was launched in April 2001 and has been promoted to the general public in both online and offline modalities. In May 2001, it was featured on hundreds of commercial Web sites in a month-long national banner-ad public service campaign through the DoubleClick advertising network, and received donated banner ad placement on the Boston Globe Web site. The site was featured as a resource on the television program CNN Presents in April 2002 (the month during which National Alcohol Screening Day occurs) [25], and has been linked from online news stories at WebMD, CNN.com, and MSNBC.com. Printed flyers (34,000) were distributed to the public through state and local health departments, alcohol treatment providers, and recovery organizations throughout the United States, and also to every employee of a major U.S. airline through May 2002. The site receives hundreds of daily referrals from Internet search engines (e.g., google.com).

Data collection

Data are collected anonymously by the Web site and cannot be traced to any identifiable individuals. Users are asked to provide their age and gender, but no further personal information is solicited. The responses to all screening questions are recorded to a secure database, as are the results of the real-time computer scoring that determines which feedback module is presented to the user. If, after viewing their screening feedback, a user immediately accesses either the "Learn More" or "Find Help" sections, their choice is recorded and associated with the screening responses and results. The user's age, gender, screening responses and results, and choice to access the "Learn More" or "Find Help" sections are linked with a record number and stored on a secure password-protected server.

The study of Web site use was approved by the Institutional Review Board at Boston University Medical Center. In addition to recording and reporting herein counts of Web site visits, page views, and completed AUDIT screening tests, we describe a sample of data entered by users between April 2, 2001, and June 2, 2002. These data were entered by users who reported they were adults 18 years or older and who completed the questionnaire. From this description, we excluded Web users who stated that they were "just curious" or that they were answering the questions for someone else.

Results are reported as descriptive statistics (means and standard deviations, proportions), and the chi-square test and t test were used as appropriate for bivariable analyses. We used analysis of variance and Duncan's multiple range test using $P = 0.05$ as the level of statistical significance to compare mean age across three drinking levels. For descriptive purposes, we defined hazardous drinking amounts consistent with NIAAA and U.S. Department of Agriculture recommendations: >14 standard U.S. drinks (13.7 g ethanol) per week for men; >7 drinks per week for women and those 65 years or older; or >4 drinks per occasion for men, >3 drinks for women. We defined those receiving an AUDIT score of ≥8 as "possible alcohol abuse or dependence," those receiving AUDIT scores <8 but exceeding consumption guidelines as "drinking hazardous amounts," and those with AUDIT scores <8 but not exceeding consumption guidelines as "nonhazardous drinkers." Drinkers with AUDIT scores of 8 or greater or drinkers of hazardous amounts when referred to together are called "hazardous drinkers."

Results

Web site use

From April 2001 to May 2003, AlcoholScreening.org received 251,170 visits (115,925/year), with the users spending an average 5 min and 25 s on the site, and yielding 109,874 completed alcohol-screening questionnaires (50,711/year). The total number of completed questionnaires as of May 16, 2004, was 180,123 in 422,324 visits.

Characteristics of Web site users

During our study period, April 2, 2001, to June 2, 2002, 66,548 users visited the site and began completing the online screening; 4,418 did not complete the screening questionnaire; 21,542 questionnaires were completed by users who stated that they were just curious about the test and related feedback or answered the questions with someone else in mind. Of the remaining 40,588 completed questionnaires, 746 users identified themselves as children under age 18 years of age or age 99 or older, and were excluded from further description or analysis. This resulted in 39,842 (35,150/year) valid questionnaires from adults (60% of adult Web site visitors) for analysis.

Of the 39,842 adults completing the screening questionnaire, the mean age was 32 (±SD 11) (range 18–98). Two-thirds of them were men, and 33% were women. Almost all users were drinking hazardous amounts (35,904/39,842, 90% [95% CI 90–90%]; 91% of men, 89% of women); 29% (95% CI 29–30%) of these hazardous drinkers (9,983/35,904) had AUDIT scores <8. More than half of the Web site visitors (21,922/39,842, 55% [95% CI 55–56%]) reported drinking hazardous amounts during a typical week. Most reported drinking binge amounts during the past year (35,181/39,842, 88% [95% CI 88–88%]; 89% of men, 87% of women).

Significantly more adults younger than 65 years old reported exceeding per occasion drinking limits compared with older adults; on the other hand, significantly more adults 65 years of age and older reported exceeding weekly consumption limits than did younger adults ($P < 0.0001$ for both comparisons) (see Table 1). Two-thirds of adults (25,921/39,842, 65% [65–66%]) had AUDIT scores of 8 or greater (possible alcohol abuse or dependence) (see Table 2). Women were significantly less likely to have scores greater than or equal to 8 (57% [95% CI 56%–58%, 7,556/13,235], compared with 69% [95% CI 68–70%, 18,365/26,607] of men, $P < 0.0001$) (see Table 3). The median AUDIT score was 10, the score at the 75th percentile was 16, and at the 90th percentile was 23.

Table 1
Age and the prevalence of hazardous per occasion and weekly amounts[a] reported by 39,842 Web site visitors completing Web-based alcohol screening

Age	Drinking hazardous per occasion (binge) amounts	Drinking hazardous weekly amounts	Total number in age group
18–24	11,346 (94)	7147 (59)	12,081
Male	7592 (94)	4550 (56)	8055
Female	3754 (93)	2597 (65)	4026
25–34	12,283 (93)	6851 (52)	13,213
Male	8652 (93)	4491 (48)	9291
Female	3631 (93)	2360 (60)	3922
35–64	11,303 (80)	7678 (54)	14,187
Male	7279 (81)	4563 (51)	8984
Female	4024 (77)	3115 (60)	5203
65 and older	249 (69)	246 (68)	361
Male	195 (70)	183 (66)	277
Female	54 (64)	63 (75)	84

Numbers in parentheses are percentages of those in the selected age group meeting the column criterion. Columns are not mutually exclusive.
[a] See text for definitions.

Table 2
Prevalence of hazardous and harmful drinking for 39,842 Web site visitors completing Web-based alcohol screening in selected age groups

Age	Nonhazardous drinkers	Drinking hazardous amounts	Possible alcohol abuse or dependence	Total number
18–24	673 (6)	2395 (20)	9013 (75)	12,081
Male	422 (5)	1347 (17)	6286 (78)	8055
Female	251 (6)	1048 (26)	2727 (68)	4026
25–34	844 (6)	3501 (27)	8868 (67)	13,213
Male	586 (6)	2159 (23)	6546 (70)	9291
Female	258 (7)	1342 (34)	2322 (59)	3922
35–64	2351 (17)	3976 (28)	7860 (55)	14,187
Male	1396 (16)	2187 (24)	5401 (60)	8984
Female	955 (18)	1789 (34)	2459 (47)	5203
65 and older	70 (19)	111 (32)	180 (50)	361
Male	55 (20)	90 (32)	132 (48)	277
Female	15 (18)	21 (25)	48 (57)	84

Numbers in parentheses are row percentages. Percentages may not add to 100 due to rounding. Columns are mutually exclusive.

Table 3
AUDIT score for 39,842 Web site visitors completing Web-based alcohol screening

AUDIT score	Male	Female	Total
0–7	8242 (31)	5679 (43)	13,921 (35)
18–24	1769 (22)	1299 (32)	
25–34	2745 (30)	1600 (41)	
35–64	3583 (40)	2744 (53)	
65 and older	145 (52)	36 (43)	
8–19	13,467 (51)	5722 (43)	19,189 (48)
18–24	4296 (53)	2002 (50)	
25–34	5150 (55)	1851 (47)	
35–64	3931 (44)	1838 (35)	
65 and older	90 (32)	31 (37)	
≥20	4898 (18)	1834 (14)	6732 (17)
18–24	1990 (25)	725 (18)	
25–34	1396 (15)	471 (12)	
35–64	1470 (16)	621 (12)	
65 and older	42 (15)	17 (20)	

Numbers in parentheses are column percentages. Columns are mutually exclusive.

Web site visitors with possible alcohol abuse or dependence were significantly younger (mean age 30.9) and more likely to be male (71%) than those drinking hazardous amounts but with AUDIT scores less than 8 (mean 33.6 years, 58% male), and nonhazardous drinkers (mean 39.0 years, 62% male) ($P < 0.0001$).

Men were more likely than women to have possible alcohol abuse or dependence, though the absolute difference was small (88% vs. 84%, OR 1.46, 95% CI 1.36–1.57). Excluding those with possible alcohol abuse or dependence (e.g., those with AUDIT scores 8 or greater), women were slightly more likely to drink hazardous amounts than men (odds ratio 1.21, 95% confidence interval 1.12–1.30).

Web site use after feedback was delivered

Almost one-fifth of the study sample (7,513/39,842, 18.9% [95% CI 18.5–19.3%]) chose the "Learn More" or "Get Help" options after receiving their results. Users with possible alcohol abuse or dependence (21.4% [95% CI 20.9–21.9%], 5,569/25,921) were more likely than users drinking hazardous amounts (15.8% [95% CI 15.1–16.5%], 1,573/9,983), who were more likely than nonhazardous drinkers (9.4% [95% CI 8.5–10.3%], 371/3,938) to choose the "Learn More" or "Get Help" options after they were presented with feedback on their drinking (Mantel-Haenszel Chi-Square test for trend $P < 0.0001$). Furthermore, of those who clicked further after getting feedback, choices were consistent with alcohol use severity: one-third (1,837/5,569, 32.9% [95% CI 31.7–34.1%]) of those with possible alcohol abuse or dependence selected "Get Help" instead of learn more, while only 8% (95% CI 6.8–9.2%, 159/1,944) of site users with AUDIT scores <8 did so ($P < 0.0001$).

Discussion

Creating and widely promoting a Web site for alcohol screening and brief intervention resulted in a significant number of Internet users visiting the site. A majority of visitors completed questionnaires and most questionnaire completers reported they were answering questions about their own drinking. Almost all reported drinking hazardous amounts, or possible alcohol abuse or dependence (AUDIT ≥ 8). What may be most remarkable about the characteristics of users of this Web site is the large proportion of women drinkers who are drinking hazardous amounts. While lower than the proportion of male hazardous drinkers in the sample, the prevalence of male hazardous drinkers was only between one and two times that of females, unlike the up to fivefold difference in hazardous drinking identified in men vs. women in primary care settings [4]. And the proportion of women drinking hazardous weekly amounts was *higher* than that for men. Use of the site was internally consistent: Web users reporting more severe problems were more likely to proceed to screens that offered more information and help. The site was feasible for screening many adults and Web site users were primarily hazardous drinkers, suggesting the possibility of reaching many people, particularly women, who otherwise might not have hazardous drinking identified or addressed.

Internet-based alcohol assessment and feedback has been reported previously. Cunningham et al. [26] posted an anonymous 21-item survey (AUDIT, typical week consumption over the past year, consequences) and materials modified from the Drinker's Check-up. Cunningham's Web address ("Try Our Free Drinking Evaluation" at http://notes.camh.net/efeed.nsf/newform) was not designed for widespread easy access though the title of the site may have garnered a number of visits. Furthermore, few of the visits generated completed questionnaires (214 of 1,729 completed questionnaires were by Web users who answered questions about their own drinking). Compared with more frequent and more consistent drinkers, less frequent drinkers and drinkers whose consumption varied a great deal over time found that the feedback given by the site was less credible. Another site, CareBetter.com, included a 43-item questionnaire followed by personalized feedback. The Web address in this case was also not transparent nor was it likely to attract users interested in evaluating their alcohol use [27]. Approximately 20% of visitors to the site completed the questionnaire. During a 172-day period, the site screened just over 2,800 individuals (approximately 10% the rate in the current study). Similar to our study, the prevalence of likely alcohol problems among completers of the questionnaires was high (89% with AUDIT ≥ 8).

AlcoholScreening.org can also be compared to another effort to screen general populations, National Alcohol Screening Day [28]. On this day implemented each year since 1999, individuals at community (e.g., hospitals and shopping malls) and primary health care sites and colleges screen volunteers. The day is widely promoted, and conducted nationwide. In community and college screenings, the AUDIT is used as the screening tool. Most of the screening activity occurs on the designated day, and almost all during the week containing that day. In 1999 at 1,089 sites, 18,043 people were screened, 43% had AUDIT scores of 8 or greater (compared with 65% of AlcoholScreening.org visitors), and 5,949 were referred for treatment. In 2002, the program screened almost 45,000 people at 2,863 sites; 12,000 were hazardous drinkers (27% vs. 90% of AlcoholScreening.org visitors) [29].

Web approaches and a national screening day rely on volunteers to seek the screening. As such, the proportion of those screened who have the target condition are relatively high compared with, for example, universal screening in general healthcare settings [30,31]. But many people screened online and at national screening events are either not having regular contact with healthcare settings or are having such contact and not being identified [27,28]. Thus Web-based screening and in-person screens are likely complementary and address problems in different populations.

Furthermore, individuals identified by in-person screening could be referred to use Web-based screening, intervention, and informational materials among other referral resources.

There are limitations to this evaluation of AlcoholScreening.org. First, because of concerns about confidentiality and our interest in seeing the feasibility of a truly anonymous screening program, we could not identify unique users of the site nor could we confirm self-reports. As a result, we may have overestimated the number of individual users because individuals may have visited the site more than once and been counted as new users. In addition, when assessing use of the Web site functions that provided more information and referral resources, we only counted users who proceeded to these parts of the site immediately after completed screening because that was the only way to be certain that the screened individual was seeking information. As a result, we may have underestimated the number of Web site users who sought help and information, as individuals may have returned during a separate Web session for this purpose. Conclusions regarding the characteristics of subjects in the sample (e.g., age, gender, alcohol use, response to feedback) should be limited to Internet users who seek and use an alcohol screening Web site. Findings may not generalize to older adults, and to populations known to use the Internet less to seek health information and less in general (e.g., lower income or education level, those without high-speed Internet connections) [14]. Finally, we were not able to assess clinical outcomes in this initial evaluation of this new technology.

Despite these limitations, we can conclude that the Web site was designed based on known valid screening approaches and included research-based intervention components, it was used extensively, and high proportions of users completed the screening. Its extensive use was likely related to concerted efforts to publicize the site, including an easily identifiable Web address recognized by common search engines.

Although screening tests are brief and valid, and professional interventions for hazardous drinking and alcohol abuse and dependence are available, many people do not avail themselves of such services [4,6]. Screening in healthcare settings is widely recommended by professional societies but screening and intervention are often not performed [30,32,33]. Some have recently suggested that screening for alcohol problems in primary care settings is inefficient and have questioned whether the effort of screening is worthwhile [31]. Most American Internet users search for health information on the Web and 9 million have searched for alcohol and drug information [14]. In a recent telephone survey, current drinkers chose computerized normative feedback more often than other options (therapist phone call, self-help book) for addressing their alcohol use [34]. Web-based assessments appear to be reliable [13]. But perhaps more importantly, they hold the promise of reaching many individuals who would otherwise receive no information or intervention. One could speculate that if these adults are similar to those seeking alcohol information at AlcoholScreening.org, 8.1 million hazardous drinkers (90% of 9 million) could be screened and receive feedback; a Web-based brief intervention, if as effective as in-person interventions, could decrease the number of hazardous drinkers by 850,000 (absolute risk reduction 10.5%) [31]. This speculation clearly goes beyond data presented herein, and points to topics for further research.

Our data suggest that Web-based screening and intervention can reach many people at low cost. Web site creation and maintenance cost $9,500, and advertising $9,000. Good search engine placement and free media coverage were the primary means of attracting visitors to the site. The volume of visitors assured high level placement in Google and other search engines. In addition, even a minimally effective intervention would have a large public health impact. Additional studies should explore the potential to reach all Internet users seeking alcohol and drug information, and should test the efficacy on drinking and other outcomes.

AlcoholScreening.org is a Web site that has a URL (Web address) with a clear message, it has been widely publicized, and it garners internally consistent responses from site visitors. The site has attracted more visitors than any other alcohol screening intervention Web site to date and it has attracted more visitors per year than a comparable in-person effort, National Alcohol Screening Day. The site has features that make Web-based screening and intervention an important addition to the public health tools available for addressing hazardous drinking: It is anonymous, accessible at any time from anywhere there is Web access, and requires minimal professional staff (and professional contact with users). Hazardous drinkers, particularly women, seek information and feedback at the site. Further research should focus on additional development of Web-based interventions, outcome evaluations using rigorous research designs, cost-effectiveness of various approaches to mass screening for alcohol problems, and research to identify the most appropriate role for Web screening (e.g., how to integrate with health professional care and other efforts such as mass in person screening). Until such results are available, Web sites such as AlcoholScreening.org can be recommended to supplement public health efforts aimed at reducing excessive alcohol use and related problems.

Acknowledgments

Join Together developed the AlcoholScreening.org Web site. Join Together is a project of Boston University supported by the Robert Wood Johnson Foundation (grant no. 2031 7). The Foundation had no role in the design, conduct, or reporting of this study. We thank Christine Lloyd Travaglini of the Data Coordinating Center, Boston University School of Public Health, for her meticulous efforts in data analysis for this report.

References

[1] The Economic Costs of Alcohol and Drug Abuse in the United States—1992. Bethesda, MD: National Institute on Drug Abuse and National Institute on Alcohol Abuse and Alcoholism, 1998 (NIH Publication no. 98-4327).

[2] Murray CJ, Lopes AD. The Global Burden of Disease. A Comprehensive Assessment of Mortality and Disability from Diseases, Injuries, and Risk Factors in 1990 and Projected to 2020. Cambridge, MA: Harvard Univ Press; 1996.

[3] Moyer A, Finney JW, Swearingen CE, Vergun P. Brief interventions for alcohol problems: a meta-analytic review of controlled investigations in treatment-seeking and non-treatment-seeking populations. Addiction 2002;97:279–92.

[4] Fiellin DA, Reid MC, O'Connor PG. Screening for alcohol problems in primary care: a systematic review. Arch Intern Med 2000; 160:1977–89.

[5] Brown RL, Saunders LA, Bobula JA, Lauster MH. Remission of alcohol disorders in primary care patients. Does diagnosis matter? J Fam Pract 2000;49:522–8.

[6] Weisner C, Matzger H. Missed opportunities in addressing drinking behavior in medical and mental health services. Alcohol Clin Exp Res 2003;27(7):1132–41.

[7] Weisner C, Mertens J, Tam T, Moore C. Factors affecting the initiation of substance abuse treatment in managed care. Addiction 2001;96:705–16.

[8] Grant BF. Barriers to alcoholism treatment: reasons for not seeking treatment in a general population sample. J Stud Alcohol 1997; 58:365–71.

[9] Hart PD, et al. "The Face of Recovery" October 2001. Survey among the recovery community conducted August 2–22, 2001 for the Alliance Project.

[10] Office of Applied Studies, Substance Abuse and Mental Health Services Administration. Health Insurance Status of Admissions for Substance Abuse Treatment: 1999. Available at http://www.samhsa.gov/oas/2k2/insuranceTX/insuranceTX.htm [accessed September 27, 2003].

[11] Saitz R, Mulvey KP, Plough A, Samet JH. Physician unawareness of serious substance abuse. Am J Drug Alcohol Abuse 1997; 23:343–54.

[12] Green-Hennessey S. Factors associated with receipt of behavioral health services among persons with substance dependence. Psychiatr Serv 2002;53:1592–8.

[13] Miller ET, Neal DJ, Roberts LJ, Baer JS, Cressler SO, Metrik J, et al. Test–retest reliability of alcohol measures: is there a difference between Internet-based assessment and traditional methods? Psychol Addict Behav 2002;16(1):56–63.

[14] Pew Internet and American Life Project. Internet Health Resources, July 16, 2003 and Internet Use by Region in the United States. Available at www.pewInternet.org [accessed February 5, 2004].

[15] Andreasen AR. Marketing Social Change: Changing Behavior to Promote Health, Social Development, and the Environment. San Francisco: Jossey-Bass Publishers; 1995.

[16] 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.

[17] Babor TF, Higgins-Biddle JC, Saunders JB, Monteiro MG. AUDIT The Alcohol Use Disorders Identification Test: Guidelines for Use in Primary Care. Second ed. Geneva, Switzerland: World Health Organization Department of Mental Health and Substance Dependence; 2001.

[18] US Department of Agriculture. Nutrition and Your Health: Dietary Guidelines for Americans. Fifth ed. Washington, DC: US Department of Health and Human Services; 2000. p. 36–7.

[19] Alcohol Research Group. 1990 National Alcohol Survey. Berkeley, CA.

[20] Bien TH, Miller WR, Tonigan JS. Brief interventions for alcohol problems: a review. Addiction 1993;88:315–36.

[21] National Institute on Alcohol Abuse and Alcoholism. The Physician's Guide to Helping Patients with Alcohol Problems. (NIH Publication no. 95-3769). Bethesda (MD): National Institute on Alcohol Abuse and Alcoholism, 1995.

[22] Saunders JB, Lee NK. Hazardous alcohol use: its delineation as a subthreshold disorder, and approaches to its diagnosis and management. Compr Psychiatry 2000;41:95–103.

[23] National Institute on Alcohol Abuse and Alcoholism. How to Cut Down on Your Drinking. (NIH Publication no. 96-3770). Bethesda (MD): National Institute on Alcohol Abuse and Alcoholism, 1996.

[24] Center for Substance Abuse Prevention. If Someone Close...has a problem with alcohol or other drugs. Available at http://www.health.org/govpubs/ph317/ [accessed September 24, 2003].

[25] Screening for Mental Health, Inc. National Alcohol Screening Day. Available at http://www.mentalhealthscreening.org/alcohol_bk.asp [accessed August 13, 2003].

[26] Cunningham JA, Humphreys K, Koski-Jannes A. Providing personalized assessment feedback for problem drinking on the Internet: a pilot project. J Stud Alcohol 2000;61:794–8.

[27] Cloud RN, Peacock PL. Internet screening and interventions for problem drinking: results from the www.carebetter.com pilot study. Alcohol Treat Q 2001;19(2):23–44.

[28] Greenfield SF, Keliher A, Sugarman D, Kozloff R, Riezes JM, Kopans B, et al. Who comes to voluntary, community-based alcohol screening Results of the first annual national alcohol screening day, 1999. Am J Psychiatry 2003;160:1677–83.

[29] NASD 5th Annual National Alcohol Screening Day. Alcohol and Health: Where do you draw the line? December 2002 eNews letter available at http://www.mentalhealthscreening.org/alcohol/enews/December2002.htm [accessed August 16, 2003].

[30] 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.

[31] Beich A, Thorsen T, Rollnick S. Screening in brief intervention trials targeting excessive drinkers in general practice: systematic review and meta-analysis. BMJ 2003;327:536–42.

[32] 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.

[33] U.S. Preventive Services Task Force. Guide to Clinical Preventive Services. Second ed. Alexandria: International Medical Publishing; 1996.

[34] Koski-Jännes, A., Cunningham J. Interest in different forms of self-help in a general population sample of drinkers. Addict Behav 2001; 26:91–9.

Primary medical care and reductions in addiction severity: a prospective cohort study

Richard Saitz[1,3,4], Nicholas J. Horton[7], Mary Jo Larson[6], Michael Winter[4,5] & Jeffrey H. Samet[1,2]

Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Department of Medicine, Boston Medical Center, Boston University School of Medicine,[1] Departments of Social and Behavioral Sciences[2] and Epidemiology,[3] Center to Prevent Alcohol Problems Among Young People,[4] Data Coordinating Center,[5] Boston University School of Public Health, Boston, MA, New England Research Institutes, Inc., Watertown, MA[6] and the Department of Mathematics, Smith College, Northampton, MA, USA[7]

Correspondence to:
Richard Saitz
Boston Medical Center
91 East Concord Street
Suite 200
Boston
MA 02118
USA
Tel: 617 414 7744
Fax: 617 414 4676
E-mail: rsaitz@bu.edu

Submitted 25 March 2004;
initial review completed 25 June 2004;
final version accepted 16 August 2004

ABSTRACT

Aims To assess whether receipt of primary medical care can lead to improved outcomes for adults with addictions.

Design We studied a prospective cohort of adults enrolled in a randomized trial to improve linkage with primary medical care.

Methods Subjects at a residential detoxification unit with alcohol, heroin or cocaine as a substance of choice, and no primary medical care were enrolled. Receipt of primary medical care was assessed over 2 years. Outcomes included (1) alcohol severity, (2) drug severity and (3) any substance use.

Findings For the 391 subjects, receipt of primary care (≥2 visits) was associated with a lower odds of drug use or alcohol intoxication (adjusted odds ratio (AOR) 0.45, 95% confidence interval (CI) 0.29–0.69, 2 d.f. χ^2 $P = 0.002$). For 248 subjects with alcohol as a substance of choice, alcohol severity was lower in those who received primary care [predicted mean Addiction Severity Index (ASI) alcohol scores for those reporting ≥2, 1 and 0 visits, respectively, 0.30, 0.26 and 0.34, $P = 0.04$]. For 300 subjects with heroin or cocaine as a substance of choice, drug severity was lower in those who received primary care (predicted mean ASI drug scores for those reporting ≥2, 1 and 0 visits, respectively, 0.13, 0.15 and 0.16, $P = 0.01$).

Conclusions Receipt of primary medical care is associated with improved addiction severity. These results support efforts to link patients with addictions to primary medical care services.

KEYWORDS Health services, primary care, severity of illness, substance abuse.

INTRODUCTION

Alcohol and drug abuse cost the United States $328 billion a year, more than heart disease or cancer [1–3]. Efficacious treatments for adults with addictions exist, yet many do not seek treatment [4,5]. Those seeking treatment may find barriers such as limited access and uncoordinated systems of care [6]. Thus, most patients who undergo detoxification do not link with addiction treatment to prevent relapse of this chronic illness [7].

There is great potential benefit to patients and providers for linking the addiction specialty treatment system with primary medical care [6]. Proven effective interventions for alcohol and drug problems can be delivered in primary care settings, such as screening, brief intervention and referral, relapse prevention and coordination of multiple specialty services (e.g. mental health, substance abuse, medical and social services) [6,8,9]; yet these systems of care typically remain unlinked and potential benefits do not accrue.

 doi:10.1111/j.1360-0443.2004.00916.x

Nevertheless, several studies have demonstrated that integrating primary medical care and addiction treatment realizes actual benefits, in particular for addiction outcomes [10,11]. While these studies of on-site care suggest benefit they are not conclusive, and furthermore they address a system of care that is currently not widely available, nor likely to be in the future given that it would require major changes in our health-care delivery systems.

Linkage of patients with addictions to primary care as it currently exists (in a distributive system of linkage/integration [12]) holds the promise of impact on addiction outcomes. Less impact might be expected from a distributive system than of a primary care intervention at an addiction treatment program, although distributive linkage, by using existing systems of care, would also probably be less costly. With tailored interventions, patients with addictions who do not have regular medical care can be linked successfully with primary care from the addiction treatment system [13]. Despite the promise of benefits, the clinical impact of this linkage model with primary care in the community remains largely unknown. As a clinical trial randomizing adults to receive or not receive primary care is unlikely to be conducted, we assessed the impact of receipt of primary medical care on addiction severity and substance use in a prospective cohort of subjects with no primary medical care.

METHODS

Subjects and design

Subjects were participants in a study of a multi-disciplinary assessment and brief motivational intervention to link adults with addictions in a residential detoxification unit, who had no primary care physician, with primary medical care. The details of the randomized trial have been published [13]. Briefly, after the acute symptoms of withdrawal had resolved, eligible subjects were enrolled and provided written informed consent. Eligible subjects were adults who spoke Spanish or English, reported alcohol, heroin or cocaine as their first or second drug of choice, and resided in proximity to the primary care clinic to which they would be referred, or were homeless. Patients with established primary care relationships they planned to continue, significant dementia, specific plans to leave the Boston area that would prevent research participation, failure to provide contact information for tracking purposes, or pregnancy were excluded. The clinical trial intervention was associated significantly with increased primary care linkage [13]. However, a substantial proportion of intervention subjects did not link with primary care (31%); and a substantial proportion of control subjects did link (53%) during the first 12 months of follow-up. Thus, the clinical trial intention-to-treat analysis was not informative regarding the impact of primary medical care on clinical outcomes. This report is not the 'treatment received' analysis of that randomized trial; rather, this prospective study focuses on the impact of exposure to what was the outcome (i.e. receipt of primary medical care) of the clinical trial and takes advantage of the prospective data collection.

Subjects were interviewed at baseline during their detoxification stay and completed up to 4 bi-annual follow-up interviews over 2 years. The current study was restricted to a prospectively enrolled and followed cohort of these trial participants who completed at least one of four scheduled research follow-up encounters, 85% (400/468) of those enrolled in the trial and alive at first follow-up opportunity. There were no statistically significant differences in subject characteristics listed below between those who entered the trial and completed versus those who did not complete any research follow-up observations except for race. Of those with follow-up, 54% were white but of those lost, 34% were white. The Institutional Review Board of Boston University 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.

Assessments

After initial resolution of the symptoms of acute withdrawal, trained research associates interviewed subjects at the detoxification unit. Assessments included: demographics, substance of choice, substances used and addiction severity [Addiction Severity Index (ASI) alcohol and drug scales] [14], substance problems [Inventory of Drug Use Consequences (InDUC-2R)] [15], readiness to change substance use [using the Stages of Change Readiness and Treatment Eagerness Scale (SOCRATES 8AOD)] [16], health-related quality of life [Short Form Health Survey (SF-36)] [17], self-report of attendance at mutual help groups such as Alcoholics Anonymous, and questions regarding primary medical care [13]. Except for demographics, all the assessments were repeated at follow-up interviews. At follow-up interviews, alcohol breath tests were performed to encourage truth telling [18].

Receipt of primary medical care was assessed during follow-up interviews using the following questions: 'Is there one particular doctor that you consider to be your regular personal doctor?'; 'Have you seen any doctors in the last 6 months (or since your last interview)?' If they did not report having a regular personal doctor but had seen a physician, they were asked: 'Would you call or go to one of these/this doctor(s) if you had a medical problem

that was not an emergency?'; 'Do you think one of these doctors could be your regular doctor?'. Subjects reporting either having or possibly having a regular personal doctor or that they would contact the doctor for non-emergent problems were asked 'What type of doctor is your regular personal/this doctor?'

While we could not assess directly the validity of self-report, we did compare self-report with administrative data sources. Computerized databases of patients seen for primary medical care at Boston Medical Center (BMC) or by Boston Health Care for the Homeless Program were queried for visits by study subjects during a 12-month period following study enrollment. This database included all visits to the two BMC-based primary care practices (>120 physicians and >70 000 visits per year), and all visits to primary health-care delivery sites for the homeless at BMC or in a city-wide network for the homeless. While subjects in the randomized trial intervention group were usually referred to care at BMC, all subjects in the cohort could pursue primary medical care anywhere. These administrative data were obtained for 95% of study subjects. Among subjects with any self-report data that were determined by administrative data to have linked, 81% (103/127) reported linkage (kappa = 0.41).

We also identified 100% of study subjects by means of substance abuse treatment utilization data obtained from the Treatment Management Information System, which is maintained by the Massachusetts Department of Public Health (DPH) Bureau of Substance Abuse Services (BSAS). These data were provided under an agreement between Boston Medical Center and the BSAS. This data set includes all episodes of substance abuse treatment utilization that occur at programs receiving state funding for addiction treatment, regardless of whether a particular treatment episode was paid for by the state. Treatment utilization was assessed for 6 months prior to study enrollment and 24 months afterwards.

Predictor variables

The main predictor variable was receipt of primary medical care. Receipt of primary care since the last research contact was defined as a visit to a primary care physician, nurse practitioner or physician assistant reported at a follow-up interview. For the visit to be defined as primary care, the subject had to report having a 'regular personal doctor', that they would call that doctor for a non-emergent issue or that they saw a doctor who 'could be their regular personal doctor'. In addition, that clinician had to be in a specialty that could be considered primary care, including obstetrics and gynecology, family medicine, pediatrics, adolescent medicine, internal medicine, AIDS doctor, asthma doctor, pulmonary doctor, cardiologist or a gastroenterologist. When the specialty was unknown to the subject or was a specialty other than those specifically queried, the physician's office was contacted to determine the specialty. Examples of specialties that were classified as non-primary care clinicians were podiatrists, emergency medicine physicians and psychiatrists. Because of its right-skewed distribution, we categorized primary care receipt for analyses as 0, 1 or 2 or more visits to avoid undue influence of outliers on the analysis. For a secondary sensitivity analysis to explore further dose–response, primary care receipt was categorized as a variable with eight levels (0, 1, 2, 3, 4, 5, 6 or 7 or more visits).

Additional predictor variables of interest included the following: demographics, homelessness (defined as one or more nights in a shelter or on the street in the preceding 6 months), addiction severity (the alcohol and drug ASI scores at study entry), attendance at mutual help meetings, the physical and mental component summary scores derived from the SF-36 [19] and the Taking Steps scale score from the SOCRATES. Substance abuse treatment utilization was a dichotomous variable of interest. This treatment variable represented any treatment, not known effective doses. The following BSAS services were considered treatment: transitional support services, recovery homes, therapeutic communities, supportive housing, residential treatment, family substance abuse shelters, day treatment, out-patient substance abuse counseling, methadone treatment, community-based case management, acupuncture, intensive out-patient treatment and postdetox recovery programs [20].

Outcome variables

Addiction severity was the primary outcome of interest. The alcohol and drug ASI composite scores (ranging from 0 to 1), and any drug use or alcohol use to intoxication (or more than three drinks in a day) (a dichotomous variable) were the main outcome variables. This latter variable was based on questions on substance use in the ASI, regarding the most recent 30 days.

A secondary analysis was conducted using substance-related problems as the dependent variable, as measured using the InDUC-2R covering a 6-month time frame (score range 0–135 [21], median of 73 observed previously for men entering drug abuse treatment [15]). Substance abuse treatment utilization (described above) was an additional outcome variable of interest, as a possible measure of impact of primary care linkage.

Subject characteristics

Of the 400 subjects eligible for this prospective cohort study, nine did not have complete data and were excluded. All subjects were interviewed at study entry: an additional 975 interviews occurred during the 2 years

after the initial interview. Of the nine subjects excluded, one was missing addiction outcome measures, five were missing addiction severity measures at study entry and three were missing homelessness, insurance and health-related quality of life data (as they completed abbreviated follow-up interviews that did not reassess these items). None were missing primary care receipt information.

Subject characteristics were similar regardless of drug of choice. Of the 391 subjects, 76% were male, mean age was 36, 50% were African American, 9% Hispanic and 60% had no health insurance; 57% of subjects with alcohol as a drug of choice were homeless, 43% of subjects with another drug as a substance of choice were homeless. The mean SF-36 Physical Component Summary (PCS) and Mental Component Summary (MCS) scores were 48 and 31, respectively (the mean MCS and PCS score for the US population is 50; 89% of adults with MCS scores of 30–34 screen positive for depression [19]). The mean ASI alcohol and drug scores were 0.47 and 0.26, respectively (these ASI scores are similar to those of individuals entering the public treatment system in Boston and more severe than those in a clinical addiction treatment sample in an HMO in California [22,23]). Subjects reporting alcohol as a drug of choice had a higher mean ASI alcohol score (0.66, SD 0.25); subjects reporting another drug as a substance of choice had a higher mean ASI drug score (0.31, SD 0.10). The mean Taking Steps scale score reflecting greater readiness to address addiction problems with higher scores was 36 (the possible range of this readiness assessment measure is 8–40 with a median of 33 for patients in alcoholism treatment [16]). Subject characteristics were similar across drug of choice groups with the exception of drug and alcohol ASI composite scores, as would be expected.

Analysis

The analysis for this paper was generated using SAS/STAT software, version 8.2 [24]. All study subjects were eligible for analyses of treatment utilization and substance use. Analyses of alcohol and drug addiction severity were restricted to subjects who reported alcohol or another drug as their substance of choice, respectively. Descriptive statistics were used to characterize the study sample (proportions, means and standard deviations). Reported *P*-values are two-tailed, and a *P*-value of less than 0.05 was considered statistically significant. Estimates of time to first report of linkage with primary medical care or substance abuse treatment, where linkage could occur at 6, 12, 18 or 24 months, were calculated using the Kaplan–Meier method. Subjects were censored after linkage or their last follow-up time.

We fitted longitudinal regression models adjusting for time and intrasubject correlations for all multi-variable analyses [25]. For dichotomous outcome variables we used generalized estimating equations (GEE) implemented in PROC GENMOD. For the continuous outcome variables we used generalized linear models for correlated data (GLMCD) implemented using PROC MIXED. Each subject contributed up to four observations of the outcome variables (from 6, 12, 18 and 24 months after study enrollment). An independent working correlation model and an empirical variance estimator were used for the GEE models, while an unstructured working covariance model was assumed for the GLMCD models. All analyses included covariates of clinical importance. Homelessness, health insurance, physical and mental health-related quality of life, readiness to change and addiction treatment, when included, were time-varying covariates. Other covariates were those assessed at study entry.

We fitted three multi-variable models to address the primary research question, one for each primary outcome. We also fitted one model for the outcome of substance abuse treatment. All analyses were adjusted for time, age, sex, race, randomized assignment in the clinical trial, homelessness, health insurance, mental and physical health-related quality of life, alcohol and drug severity at study entry and readiness to change. Analyses with addiction severity and substance use as the outcome variables were also adjusted for substance abuse treatment and use of mutual help groups. Analyses with substance abuse treatment and substance use as outcomes were adjusted for substance of choice.

Secondary confirmatory/sensitivity analyses addressed whether the results of the three main models and the substance abuse treatment model would differ when entering primary care as an eight-level variable. An additional analysis confirmed the results by testing the effect of primary care receipt on substance-related problems, adjusting for the same covariates as in the models with substance use as the outcome. To address the possibility of higher-order effects, we added the following seven interactions with receipt of primary medical care to the three main models: alcohol and drug severity at study entry, mutual help use, readiness to change, physical health-related quality of life, randomization assignment and substance abuse treatment utilization.

RESULTS

Of 391 subjects, 194 (49.6%) received a total of two or more primary care visits, 53 (13.5%) received a total of one visit and 144 (36.8%) received no primary medical care visits during the study period. Kaplan–Meier estimated proportions of subjects receiving a total of one or more primary care visits at 6, 12, 18 and 24 months after study entry were 31%, 44%, 58% and 70%, respectively.

The four corresponding proportions for receipt of a total of two or more primary care visits were 19%, 30%, 44% and 58%.

Receipt of primary medical care and addiction severity and substance abuse treatment

For subjects who reported alcohol as their first or second drug of choice, receipt of primary care was significantly associated with improved alcohol (ASI) severity ($F_{2,239} = 3.29$, $P = 0.04$) (predicted mean alcohol ASI for 0, 1, ≥2 visits, respectively, 0.34, 0.26, 0.30) in a multivariable analysis. Similarly, for subjects reporting heroin or cocaine as their first or second drug of choice, receipt of primary care was associated significantly with improved drug (ASI) severity ($F_{2,291} = 4.49$, $P = 0.01$) (predicted mean drug ASI for 0, 1, ≥2 visits, respectively, 0.16, 0.15, 0.13). For all 391 subjects, regardless of substance of choice, receipt of primary care was associated significantly with a decreased odds of drug use or use of alcohol to intoxication (during 30 days) ($\chi^2 = 12.90$, 2 d.f., $P = 0.0016$). During the 2 years of follow-up, 267 subjects (67.3%) received substance abuse treatment services (Kaplan–Meier estimates for each of the four successive 6-month periods 33.5%, 44.3%, 59.9%, 73.5%). In a multi-variable analysis, receipt of primary care was not significantly associated with receipt of substance abuse treatment (Table 1).

Confirmatory/sensitivity analyses

In multi-variable models entering the same covariates as the previously described models, an eight-level variable for receipt of primary care was associated with lower alcohol severity at a borderline level of significance (decrease in alcohol severity score for additional visit −0.009, 95% CI −0.020–0.001, $F_{1,239} = 2.88$, $P = 0.09$), was significantly associated with improved drug severity (decrease in drug severity for additional visit, −0.006, 95% CI −0.010 to −0.001, $F_{1,290} = 6.70$, $P = 0.01$), and was significantly associated with a reduced odds of drug use or drinking alcohol to intoxication (OR for one more visit 0.83, 95% CI 0.75–0.91, $\chi^2 = 15.76$, 1 d.f., $P < 0.0001$). This eight-level ordinal variable was not sig-

Table 1 Association between primary care visits and addiction outcomes in multi-variable analyses.

	Substance abuse treatment n = 391 Odds ratio, 95% CI	*Alcohol severity n = 248*	*Drug severity n = 300*	*30-day drug use or use of alcohol to intoxication n = 391 Odds ratio, 95% CI*
Primary care visits during 6 months		Predicted mean ASI score	Predicted mean ASI score	
0	–	0.34	0.16	–
1	1.08 (0.70–1.67)	0.26	0.15	0.91 (0.54–1.52)
≥ 2	1.04 (0.73–1.49)	0.30	0.13	0.45 (0.29–0.69)
P-value	$P = 0.94$	$P = 0.04$	$P = 0.01$	$P = 0.002$
		Parameter estimate (95% CI)	Parameter estimate (95% CI)	
Age (decade)	0.70 (0.58–0.86)	−0.0205 (−0.0502, 0.0132)	−0.0014 (−0.0161, 0.0134)	0.86 (0.65–1.15)
Homeless	1.52 (1.09–2.12)	0.0448 (0.0023, 0.0872)	0.0173 (−0.0002, 0.0349)	2.52 (1.72–3.69)
Health insurance	1.73 (1.21–2.47)	0.0392 (−0.0068, 0.0851)	0.0152 (−0.0022, 0.0325)	1.86 (1.28–2.72)
Mental health (MCS) (10 points)	0.88 (0.79–0.99)	−0.0636 (−0.0792, −0.0479)	−0.0319 (−0.0381, −0.0258)	0.67 (0.58–0.77)
Physical health (PCS) (10 points)	1.20 (1.02–1.40)	−0.0392 (−0.0590, −0.0194)	−0.0192 (−0.0273, −0.0112)	0.85 (0.70–1.02)
Baseline addiction severity, alcohol (0.10 points)	0.95 (0.90–1.01)	0.0176 (0.0076, 0.0276)	−0.0046 (−0.0077, −0.0015)	1.02 (0.94–1.09)
Baseline addiction severity, drug (0.10 points)	1.06 (0.93–1.20)	−0.0117 (−0.0283, 0.0049)	0.0234 (0.0136, 0.0331)	1.14 (0.98–1.34)
Taking Steps score (1 point)	1.05 (1.03–1.08)	−0.0088 (−0.0121, −0.0054)	−0.0026 (−0.0040, −0.0012)	0.86 (0.83–0.89)

For primary care visits, reference group for odds ratios is no primary care visits. All analyses are adjusted for the variables listed in the table and sex, race/ethnicity and randomized group. Time (1st, 2nd, 3rd or 4th interview) was a significant predictor of outcome only for the substance abuse treatment analysis ($\chi^2 = 29.53$, d.f. = 3, $P < 0.0001$). ASI and substance use outcome analyses also adjusted for substance abuse treatment and mutual help use (not significant in any analysis). Substance use and substance abuse treatment outcome analyses adjusted for substance of choice (not statistically significant).

nificantly associated with receipt of substance abuse treatment (OR for an additional visit 1.00, 95% CI, 0.93–1.08, $\chi^2 = < 0.01$, 1 d.f., $P = 0.97$).

For all 391 subjects, regardless of substance of choice, receipt of primary care (categorized as 0, 1 or ≥2 visits) was significantly associated with decreased substance-related problems (measured by the InDUC-2R, a measure of problems in past 6 months or since last research assessment if >6 months) ($F_{2,381} = 13.59$, $P < 0.0001$) (predicted mean score for 0, 1, ≥2 visits, respectively, 52, 51, 39).

To address the possibility of higher order effects, we tested seven interactions between receipt of primary care and alcohol and drug severity at study entry, mutual help use, the Taking Steps score, physical health-related quality of life, randomization group and substance abuse treatment utilization in models with alcohol and drug severity and drug use or alcohol intoxication outcomes (three outcomes, three models). Of these 21 interactions tested, only one was significant at $P < 0.05$, the interaction with physical health-related quality of life in the model predicting drug use severity (Table 2) ($F_{2,291} = 3.53$, $p = 0.03$). The association between primary care receipt and drug use severity varied by physical health-related quality of life. Predicted mean differences in drug severity were similar regardless of primary care receipt for those with better health (higher PCS). Primary care appeared to have a greater impact on drug severity in people with worse physical health status (lower PCS scores), than it did on people with better health status. At the observed mean PCS of 48, and at 2 standard errors above this mean (70), there was little difference in the decreases in ASI drug score attributable to receipt of primary care.

CONCLUSIONS

Receipt of primary medical care in a distributive model by adults with addictions who have not recently had such care is associated with reduced problems and severity of addictions over a 24-month period. This association does not appear to be mediated by exposure to substance abuse treatment, nor does it appear to be affected by addiction severity, mutual help use or readiness to change. As might be expected, primary care had a greater impact on addiction severity in patients with worse physical health. However, given the numerous interactions tested and the appearance of the finding for subjects with only cocaine or heroin as a substance of choice, these latter findings should be viewed as hypothesis generating. Analyses suggest some evidence for dose–response, both those that tested levels of primary care exposure and the ordered results from all other models (i.e. more effect from ≥2 visits versus 1, and 1 versus 0 visits). In the analysis of alcohol ASI results were not ordered according to primary care dose, but the one-visit group was substantially smaller than the 0- or ≥2-visit groups.

Our findings are consistent with prior studies. In a landmark randomized trial, substance abuse treatment patients receiving on-site medical, psychiatric, employment and family services had less opiate use and improved medical, employment, legal and psychiatric outcomes [26]. In another randomized trial patients with substance abuse-related medical conditions assigned to on-site primary care at an addiction treatment program were more likely to remain abstinent than patients in usual separate primary care [10]. In a retrospective cohort study patients with addictions treated at programs with on-site primary care had improved addiction outcomes when compared with patients treated in programs without such services [11]. Willenbring *et al.* tested another variation of on-site medical and addictions care integration [27]. In a randomized trial in a special alcohol clinic for veterans, the integrated care group was more likely to be abstinent than a usual care group [28]. Laine *et al.* found that patients with addictions who receive both regular addiction and medical (not necessarily primary) care are less likely to be hospitalized than those who received one or neither service [29]. Thus, medical care in general, and primary care specifically, appears to improve addiction outcomes.

Our study adds to this literature by suggesting that these results are also true for primary medical care as it is delivered in the community to a group of patients with addictions that include many without homes or

Table 2 Predicted mean differences in drug addiction severity by receipt of primary care for three values of physical health-related quality of life (sample mean 48, and 2 standard errors above and below).

	Predicted mean differences in drug addiction severity		
	Sample mean - 2 SE	*Sample mean*	*Sample mean + 2 SE*
Receipt of primary care	PCS = 26	PCS = 48	PCS = 70
No primary care visits	−0.08	−0.15	−0.21
1 visit	−0.19	−0.18	−0.19
≥ 2 visits	−0.16	−0.19	−0.23

health insurance. In fact, it is people who are socially disadvantaged who may be most likely to benefit from primary medical care. Laine *et al.*'s population, who received benefit from regular care, were insured by Medicaid (state health insurance for the poor or disabled) [29]. Gelberg *et al.* have noted previously that in a vulnerable population, more severe addictions and more severe homelessness did not deter access to needed care, and that having a regular source of care predicted better health outcomes [30]. The observation that substance abuse treatment did not appear to mediate the relationship between primary care and improved addiction severity is not surprising. Linkage between primary care and the addiction treatment system has been elusive [6]. Most patients with addictions do not seek specialty treatment [5] and substance abuse treatment has been notably not associated with linkage to primary medical care [31]. This absence of apparent contribution of substance abuse treatment to the observed benefit of primary care to addictions outcomes could be viewed as both a cup half empty and half full. Primary care may be contributing independently to improved addictions outcome. In addition, the opportunity to achieve even greater benefit may exist if mechanisms were instituted to link primary care more effectively with substance abuse treatment.

The principal limitation that should be considered in interpreting our results is that our data are from an observational study. As a result, the observed associations between receipt of primary care and addiction severity could be due to confounding. However, we adjusted for likely confounders of this relation. Furthermore, the fact that data for this observational study were collected prospectively in a study focused on primary care and addictions is a particular strength. One might also question whether primary care led to reduced addiction severity, or whether adults with improved addictions became health conscious and linked with primary care. Another similar possibility is that those with the most severe problems who 'hit rock bottom' began to take care of themselves, including a visit to primary care and abstinence or less use. Both explanations seem less plausible because analyses adjusted for substance abuse consequences and health status. Furthermore, measures of addiction severity referred to the recent past while measures of primary care receipt referred to the past 6 months or more, making the former more likely. The temporal association, the confirmation of a prior hypothesis, adjustment for important possible confounders and the consistency across outcomes suggest that receipt of primary care was associated with later improvements in addiction severity. Finally, some of the observed effects were relatively small (e.g. differences of 0.02 in drug ASI). Our goal was not to determine clinical significance but rather to identify an association. It is notable that an association between primary medical care and addiction severity was detectable at all, given the variability in subject characteristics and the many other contributors to addiction severity. In fact, some effects were quite substantial and clinically significant (e.g. >50% decrease in odds of substance use, large changes in InDUC problem score and alcohol ASI).

The important question, of whether or not primary medical care as it is currently delivered in the community can improve addiction severity, is best answered in a naturalistic study, and this highlights several important strengths of this study. First, the data collection in this cohort study was prospective. Secondly, two of the main purposes of the research assessments a priori were to assess primary medical care utilization and addiction severity in detail. Thirdly, multiple measures of addiction severity used in the study have been validated and results across these measures were consistent. Finally, receipt of primary care was corroborated by administrative data.

One could still argue that our findings are due to inadequately addressed confounding or that the outcome (i.e. addiction severity) leads to seeking primary medical care. Such a critical assessment might contend that the hypothesis, 'primary medical care leads to improved addiction severity', can only be answered definitively by a randomized trial of primary care. Addressing these concerns optimally is difficult, as such a study will probably not be conducted, for pragmatic and ethical reasons.

This study and previous studies support the contention that receipt of primary care is associated with improved outcomes for adults with addictions. Our study results and those of others cited herein may be the best type of evidence to bring to address the question. This evidence suggests strongly that patients with addictions should receive primary medical care in addition to addiction treatment. Efforts to link and/or integrate primary medical care with addiction specialty care are worthwhile. Patients with addictions can be linked with primary medical care [13]. In addition, actually providing services known to be efficacious and feasible in primary care settings that are not currently in widespread use could lead to even greater improvements for patients with addictions. How to make primary medical care truly accessible and actually used by adults with alcohol and other drug dependence remains a challenge for researchers and clinicians alike.

Acknowledgements

Preliminary results of this study were presented at the annual national meeting of the Association for Medical

Education and Research in Substance Abuse (AMERSA). November 2003, in Baltimore, MD. Data management was performed by the Data Coordinating Center, Boston University School of Public Health. The National Institute on Alcohol Abuse and Alcoholism (RO1-AA10870), and the National Institute on Drug Abuse (RO1-DA10019) provided support for the study. The Bureau of Substance Abuse Services, Department of Public Health, Commonwealth of Massachusetts, provided substance abuse treatment utilization data, and the Boston Health Care for the Homeless Program provided data on primary care utilization, as did Boston Medical Center. This work was supported in part by General Clinical Research Center grant MO1-RR00533 from the National Center for Research Resources.

References

1. Harwood, H. J., Fountain, D. & Livermore, G., eds (1998) *The Economic Costs of Alcohol and Drug Abuse in the United States—1992*. Washington, DC: National Institute on Alcohol Abuse and Alcoholism.
2. Anonymous (2003) *NHLBI Morbidity and Mortality Chartbook, 2000*. Bethesda, MD: National Heart, Lung, and Blood Institute.
3. Harwood, H. (2000) *Updating Estimates of the Economic Costs of Alcohol Abuse in the United States: Estimates, Update Methods and Data*. Bethesda, MD: The Lewin Group for the National Institute on Alcohol Abuse and Alcoholism.
4. Miller, W. R. & Wilbourne, P. L. (2002) Mesa Grande: a methodological analysis of clinical trials of treatments for alcohol use disorders. *Addiction*, 97, 265–277.
5. Green-Hennessy, S. (2002) Factors associated with receipt of behavioral health services among persons with substance dependence. *Psychiatric Services*, 53, 1592–1598.
6. Samet, J. H., Friedmann, P. & Saitz, R. (2001) Benefits of linking primary medical care and substance abuse services: patient, provider, and societal perspectives. *Archives of International Medicine*, 161, 85–91.
7. Mark, T., Dilonardo, J., Chalk, M. & Coffey, R. (2002) *Substance Abuse Detoxification: Improvements Needed in Linkage to Treatment (SMA-02-3728)*. Rockville, MD: Substance Abuse and Mental Health Services Administration.
8. Friedmann, P. D., Saitz, R. & Samet, J. H. (1998) Management of adults recovering from alcohol or other drug problems: relapse prevention in primary care. *JAMA*, 279, 1227–1231.
9. Moyer, A., Finney, J. W., Swearingen, C. E. & Vergun, P. (2002) Brief interventions for alcohol problems: a meta-analytic review of controlled investigations in treatment-seeking and non-treatment-seeking populations. *Addiction*, 97, 279–292.
10. Weisner, C., Mertens, J., Parthasarathy, S., Moore, C. & Lu, Y. (2001) Integrating primary medical care with addiction treatment: a randomized controlled trial. *JAMA*, 286, 1715–1723.
11. Friedmann, P., Zhang, Z., Hendrickson, J., Stein, M. & Gerstein, D. (2003) Effects of primary medical care on addiction and medical severity in substance abuse treatment programs. *Journal of General Internal Medicine*, 18, 1–8.
12. Samet, J. H., Saitz, R. & Larson, M. J. (1996) A case for enhanced linkage of substance abusers to primary medical care. *Substance Abuse*, 17, 181–192.
13. Samet, J. H., Larson, M., Horton, N. J., Doyle, K., Winter, M. & Saitz, R. (2003) Linking alcohol and drug dependent adults to primary medical care: a randomized controlled trial of a multidisciplinary health intervention in a detoxification unit. *Addiction*, 98, 509–516.
14. McLellan, A. T., Kushner, H., Metzger, D., Peters, R., Smith, I., Grissom, G., Pettinati, H. & Argeriou, M. (1992) The fifth edition of the Addiction Severity Index. *Journal of Substance Abuse Treatment*, 9, 199–213.
15. Tonigan, J. & Miller, W. (2002) The inventory of drug use consequences (InDUC): test–retest stability and sensitivity to detect change. *Journal of Psychological Addictive Behaviors*, 16, 165–168.
16. Miller, W. & Tonigan, J. (1996) Assessing drinkers' motivation for change: the Stages of Change Readiness and Treatment Eagerness Scale (SOCRATES). *Journal of Psychological Addictive Behaviors*, 10, 81–89.
17. Ware, J. E. (1993) *SF-36 Health Survey Manual and Interpretation Guide*. Boston, MA: The Health Institute.
18. Welte, J. W., Perry, P., Longabaugh, R. & Clifford, P. R. (1998) An outcome evaluation of a hospital-based early intervention program. *Addiction*, 93, 573–581.
19. Ware, J. E. (1994) *SF-36 Physical and Mental Health Summary Scales: a User's Manual*. Boston, MA: The Health Institute.
20. Cellucci, A. P., Swift, J., O'Leary, W. D., Koh, H. K. & Walker, D. K. (2001) *Substance Abuse Prevention and Treatment Service Directory*. Boston, MA: Massachusetts Department of Public Health Bureau of Substance Abuse Services.
21. Miller, W. & Tonigan, J. (1995) *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.
22. Leonhard, C., Mulvey, K., Gastfriend, D. R. & Shwartz, M. (2000) The Addiction Severity Index: a field study of internal consistency and validity. *Journal of Substance Abuse Treatment*, 18, 129–135.
23. Weisner, C., McLellan, A. T. & Hunkeler, E. M. (2000) Addiction severity index data from general membership and treatment samples of HMO members. One case of norming the ASI. *Journal of Substance Abuse Treatment*, 19, 103–109.
24. SAS Institute (1999–2001) *SAS/STAT software*, version 8.2, *SAS System for Windows*. Cary, NC: SAS Institute, Inc.
25. Cnaan, A., Laird, N. M. & Slasor, P. (1997) Using the general linear mixed model to analyse unbalanced repeated measures and longitudinal data. *Statistics in Medicine*, 16, 2349–2380.
26. McLellan, A. T., Arndt, I. O., Metzger, D. S., Woody, G. E. & O'Brien, C. P. (1993) The effects of psychosocial services in substance abuse treatment. *JAMA*, 269, 1953–1959.
27. Willenbring, M. L., Olson, D. H. & Bielinski, J. (1995) Integrated out-patients treatment for medically ill alcoholic men: results from a quasi-experimental study. *Journal of Studies on Alcohol*, 56, 337–343.
28. Willenbring, M. L. & Olson, D. H. (1999) A randomized trial of integrated out-patient treatment for medically ill alcoholic men. *Archives of Internal Medicine*, 159, 1946–1952.
29. Laine, C., Hauck, W. W., Gourevitch, M. N., Rothman, J., Cohen, A. & Turner, B. J. (2001) Regular out-patient

medical and drug abuse care and subsequent hospitalization of persons who use illicit drugs. *JAMA*, **285**, 2355–2362.

30. Gelberg, L., Andersen, R. M. & Leake, B. D. (2000) The behavioral model for vulnerable populations: application to medical care use and outcomes for homeless people. *Health Services Research*, **34**, 1273–1302.
31. Saitz, R., Mulvey, K. P. & Samet, J. H. (1997) The substance-abusing patient and primary care: linkage via the addiction treatment system? *Substance Abuse*, **18**, 187–195.

Linkage with Primary Medical Care in a Prospective Cohort of Adults with Addictions in Inpatient Detoxification: Room for Improvement

Richard Saitz, Mary Jo Larson, Nicholas J. Horton, Michael Winter, and Jeffrey H. Samet

Objective. To identify patient characteristics and health care experiences associated with primary care linkage after alcohol or drug detoxification.

Data Sources/Study Setting. Primary data collected over two years. Subjects were adults without primary medical care, in an urban residential detoxification program.

Study Design. A prospective cohort study in the context of a randomized trial of a linkage intervention, and an expansion of Medicaid benefits.

Data Collection/Extraction Methods. Data were collected by interview assessment of predisposing, enabling, and illness variables. Linkage was defined as self-report of at least one visit with a primary care clinician during follow-up.

Principal Findings. Of 400 subjects, 63 percent linked with primary medical care. In a multivariable model adjusting for randomization assignment, predisposing, enabling, and illness variables, women, those with no recent incarceration, those with support for abstinence by family or friends, and those who had visited a medical clinic or physician recently were significantly more likely to link with primary care. Those with health insurance during follow-up were also more likely to link. Recent mental health or addictions treatment utilization and health status were not associated with linkage.

Conclusions. A substantial proportion of adults with addictions do not link with primary medical care. These data suggest that efforts could be focused on those least likely to link, that contacts with mental health and addictions treatment providers are underutilized opportunities for these efforts, and that health policy changes such as expanding health insurance benefits may improve entry of substance-dependent patients into primary medical care.

Key Words. Primary care, addictions, health insurance, cohort, linkage

Many patients with addictions do not receive primary medical care. For example, of persons entering addiction treatment in Boston, only 41 percent reported having a primary care physician (Saitz, Mulvey, and Samet 1997). Yet these patients have many acute and chronic medical illnesses, both related and unrelated to their addictions (DeAlba, Samet, and Saitz in press; Saitz

2003). They also tend to use more costly episodic care for medical needs, such as the emergency department (McGeary et al. 2000; French et al. 2000). To address this shortfall there have been calls to link substance abuse treatment with primary care (Morris 1995; Levin et al. 1993; Samet, Friedmann, Saitz 2001). But how to link these systems to benefit patients, and which patients with addictions are at greatest risk of going without primary care, is unknown. In fact, several federal agencies have recently attempted to gather expertise and develop research agendas to answer these questions and have ongoing requests for research in this area (National Institute of Mental Health 2003).

Primary care can lead to better health for many groups of patients (Starfield 1998). In a recent retrospective cohort study, adults with addiction who received regular primary medical care were less likely to be hospitalized (Laine et al. 2001). In another study, onsite primary care was associated with improved addiction severity (Friedmann et al. 2003). In a randomized clinical trial, adults with addiction and substance-abuse-related medical conditions randomized to receive primary medical care integrated with their addictions care were more likely to be abstinent (Weisner et al. 2001). These studies confirm some of the predicted benefits of linkage of persons with addictions to primary medical care (Levin et al. 1993; Schlenger et al. 1992; Samet, Saitz, and Larson 1996). Recently, new therapeutic options have been proven effective in primary care settings, such as naltrexone for alcoholism, and buprenorphine for office-based opioid dependence treatment (O'Connor

---

Preliminary results were presented at the annual national meetings of the Society of General Internal Medicine (SGIM) in May 2000 in Boston, at the College on Problems of Drug Dependence (CPDD) in June 2000 in San Juan, Puerto Rico, and the Association for Medical Education and Research in Substance Abuse (AMERSA) in November 2000 in Alexandria, VA.

This work was supported by a grant from the National Institute on Alcohol Abuse and Alcoholism (R01 AA10870) and the National Institute on Drug Abuse (R01-DA10019). Dr. Saitz was a Generalist Physician Faculty Scholar of the Robert Wood Johnson Foundation during the time this study was done. This work was supported in part by General Clinical Research Center grant no. MO1 RR00533 from the National Center for Research Resources.

Richard Saitz, M.D., M.P.H., and Jeffrey H. Samet, M.D., M.A., M.P.H., are from the Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Department of Medicine, Boston Medical Center and Boston University School of Medicine. Additionally, Dr. Saitz is with the Department of Epidemiology and Dr. Samet is with the Department of Social and Behavioral Sciences, Boston University School of Public Health. Michael Winter, M.P.H., is with the Data Coordinating Center, Boston University School of Public Health. Mary Jo Larson, Ph.D., is with New England Research Institutes, Watertown, MA. Nicholas J. Horton, Sc.D., is with the Department of Mathematics, Smith College, Northampton, MA.

et al. 1998; O'Malley 2003). Other theoretical benefits are more efficient use of health services, opportunity for preventive health interventions, and better health.

Some addiction treatment programs offer primary medical care onsite (Weisner et al. 2001). One study of integrated medical care for patients with alcoholism has even shown a mortality benefit (Willenbring and Olson 1999). But these integrated programs require significant changes in existing systems. The more common pattern of organization of primary care and specialty addiction treatment services in the United States is no relationship, or a distributive one, rather than an integrative, onsite model. In the distributive pattern, medical and addictions care are delivered in separate locations, and patients and information are transferred from one location to another (Samet, Saitz, and Larson 1996). Yet this flow is anything but seamless (Samet, Friedmann, and Saitz 2001), since receipt of addictions care is not always associated with linkage to primary medical care (Saitz, Mulvey, and Samet 1997). In substance abuse programs opportunities to link patients with primary care are being missed.

Because regular medical care has been shown to have benefits, the need to improve linkage of addicted persons with primary medical care has become more urgent. A recent randomized trial found that an onsite multidisciplinary health evaluation, including motivational counseling, could increase linkage to primary medical care for adults with addictions and no regular physician (Samet et al. 2003). A greater understanding of which patients with addictions are more or less likely to link with medical care would help target such interventions and help in the design of additional efforts to link patients with primary medical care. In a retrospective cross-sectional study of patients entering addiction treatment, those more likely to have primary medical care were older, female, had health insurance, and had medical illness (Saitz, Mulvey, and Samet 1997). The study was limited, however, due to an inability to distinguish the temporal relationship between risk factors and linkage and to characterize those who link to medical care after contact with the addiction treatment system. Factors associated with linkage to primary medical care may certainly be different in patients with addictions undergoing detoxification than they are for the general population or for patients with addictions who are not yet receiving any specialty care. And patients undergoing detoxification, many of whom contact the health care system only at detoxification programs and emergency departments, are reachable and could potentially be connected with needed primary medical care, particularly if these efforts could be focused on those less likely to link without them.

Therefore, the objective of this study was to examine, using a prospective design, characteristics and health care experiences of adults with addictions associated with linkage to primary medical care. We hypothesized that greater addiction severity and access barriers (e.g., lack of insurance, ethnicity, incarceration) would interfere with linkage to primary medical care (Saitz, Mulvey, and Samet 1997; McGeary et al. 2000; French et al. 2000; Conklin, Lincoln, and Tuthill 2000; Hargraves, Cunningham, and Hughes 2001). We also hypothesized that women, those recognizing their substance problem and having social support for recovery, those believing medical care was important, those with worse health status, and those who had prior contacts with addiction, episodic medical or mental health specialty care, would be more likely to link with primary medical care (Saitz, Mulvey, and Samet 1997; McCarthy et al. 2002; Diamant et al. 2001).

METHODS

Design

The study was a prospective cohort. It was approved by the Institutional Review Board at Boston University Medical Center, and additional confidentiality protection was provided by a certificate of confidentiality provided by the National Institute on Alcohol Abuse and Alcoholism. All subjects provided written informed consent.

Subjects

All 2,062 adults admitted to and voluntarily staying at a free-standing urban residential alcohol and drug detoxification unit between June 1, 1997, and April 1, 1999, were screened for the study when research staff and patients were both available (Samet et al. 2003). They were screened for eligibility and enrolled on their second day or later in the detoxification unit. Inclusion criteria were the following: (1) alcohol, heroin, or cocaine as the patient's first or second drug of choice; (2) age greater than 17 years; and (3) residence in proximity to a referral primary care clinic or homelessness. The exclusion criteria were as follows: (1) an established primary care relationship that the patient intended to continue (980 persons, 69 percent of those ineligible); (2) mental deficiencies making the subject unable to provide pertinent history or informed consent (score of less than 21 of 30 on the Mini-Mental State Examination) (Folstein, Folstein, and McHugh 1975); (3) specific plans to leave the area in the next 12 months; (4) inability to provide three contact

names for follow-up tracking; (5) pregnancy; and (6) not fluent in English or Spanish. Of 642 eligible subjects, 470 provided consent and were enrolled in the cohort. All 470 participated in the randomized clinical trial, the Health Evaluation and Linkage to Primary care (HELP) study (Samet et al. 2003). All subjects were randomly assigned to receive either standard medical care referral by clinical addictions treatment staff on an as needed basis (usual care), or an enhanced effort for referral to primary medical care. This effort involved assessment by a physician, nurse, and social worker to address medical, psychological, and social issues, brief counseling by these providers trained in motivational interviewing to encourage primary care linkage, and making a specific appointment with a primary care physician and letters and phone calls to facilitate linkage (Samet et al. 2003). Neither option included ongoing primary medical care at the detoxification unit.

Assessments

After initial resolution of the symptoms of acute withdrawal during the first 24 hours, subjects were interviewed at the detoxification unit by trained research associates. Assessments included demographics, health care utilization, social support, barriers to primary care linkage, beliefs about primary care, substances used, addiction severity (Addiction Severity Index [ASI] alcohol, drug, and psychological sub-scales) (McLellan et al. 1992), consequences of drug use (Inventory of Drug Use Consequences [INDUC-2L]) (Miller and Tonigan 1995), readiness to change substance use (using the Stages of Change Readiness and Treatment Eagerness Scale, [SOCRATES 8AOD]) (Miller and Tonigan 1996), depressive symptoms (Center for Epidemiologic Studies Depression [CES-D] scale) (Radloff 1977), health-related quality of life (Short Form Health Survey [SF-36]) (Ware 1993), and questions regarding comorbid medical diagnoses (Stein et al. 1998; Saitz, Mulvey, and Samet 1997).

Primary care linkage was determined by asking: "Is there one particular doctor that you consider to be your regular personal doctor?"; "Have you seen any doctors in the last six months (or since your last interview)?" If they did not report having a regular personal doctor but had seen a physician, they were asked: "Would you call or go to one of these/this doctor(s) if you had a medical problem that was not an emergency?"; "Do you think one of these doctors could be your regular doctor?" Subjects reporting either having or possibly having a regular personal doctor or that they would contact the doctor for nonemergent problems were asked, "What type of doctor is your regular personal/this doctor?"

For these subjects being detoxified from substances, problem use of alcohol or other drugs was defined as either frequent use ($\geq$ 3 times per week) for a year or more, or 5 or more days of use in the past 30 days for any substances listed in the ASI. For this problem use definition, alcohol use was defined as either intoxication or three or more drinks on one day (McLellan et al. 1992; Volpicelli et al. 1992).

Outcomes

Outcomes were assessed by in-person interview (phone as a secondary option) at 6, 12, 18, and 24 months after baseline. Time to first self-reported linkage to primary medical care during the 24 months following study enrollment was the primary outcome of this cohort study (as well as of the randomized trial), where first linkage could occur at 6, 12, 18, or 24 months. Linkage to primary care was defined as at least one visit to a primary care physician, nurse practitioner, or physician assistant. For the visit to be defined as primary care, the subject had to report having a "regular personal doctor," that they would call this doctor for a nonemergent issue, or that they saw a doctor that "could be their regular personal doctor." The clinician had to be in a specialty that could be considered primary care, including obstetrics and gynecology, family medicine, pediatrics, adolescent medicine, internal medicine, AIDS doctor, asthma doctor, pulmonary doctor, cardiologist, or a gastroenterologist. When the specialty was unknown to the subject or was a specialty other than those specifically queried, the physician's office was contacted to determine the specialty.

While we could not directly assess the validity of self-report, we did compare self-report with administrative data sources. Computerized databases of patients seen for primary medical care at Boston Medical Center (BMC) or by Boston Health Care for the Homeless Program were queried for visits by study subjects during a 12-month period following study enrollment. This database included visits to the two BMC-based primary care practices ($>$120 physicians and $>$50,000 visits per year), and visits to primary health care delivery sites for the homeless at BMC or in a citywide network for the homeless. Whereas subjects in the randomized trial intervention group were usually referred to care at BMC, all subjects in the cohort could pursue primary medical care anywhere. Administrative data were obtained for 95 percent of study subjects. Among subjects with any self-report data that were determined by administrative data to have linked, 81 percent (103/127) reported linkage (Kappa = 0.41).

Independent Variables and Statistical Analysis

Analyses used survival methods with time to event defined as the number of months between randomization and report of primary care linkage over the 24-month follow-up period. To aid in understanding the rate of linkage (accounting for censoring after linkage or last follow-up), we calculated annualized rates of linkage by dividing the number of linkage events by person-years of follow-up. Initial review of predictors of linkage was undertaken using stratified bivariable analysis.

Multivariable proportional hazards regression models (Cox) were used to compare the hazard ratio for different predictor variables while accounting for other factors. We used the original theoretical framework provided by Andersen et al. to understand determinants of medical care utilization in the United States as guidance for analyses, as well as Gelberg et al.'s extension of this framework to include domains for vulnerable populations (Andersen 1995; Gelberg, Andersen, and Leake 2000; Andersen and Newman 1973). Specific vulnerable population variables were incarceration, perception of medical need ("How important to you is treatment for medical problems," from the ASI), substance abuse problem recognition (SOCRATES), and mental illness symptoms (CES-D, suicide attempt history, and psychiatric medication use). We also included one system variable: randomized group. We chose variables for entry into multivariable models based on review of the literature, clinical importance, bivariable analyses (using the log-rank test to determine statistical significance) using the liberal criterion $p<0.20$, and attention to conceptual overlap that could lead to collinearity (for example, health-related quality of life and report of a chronic medical illness). We then constructed multivariable models sequentially, in the order consistent with the theoretical framework. All models included age, gender, race, and randomization assignment. The first model also included predisposing variables. The second model added enabling variables to those variables found to be significant at $p<.20$ in the first model. To assess the effect of a variable not collected at baseline, the third model added health insurance as a time-variable predictor (same six-month time period as the report of linkage) to variables significant ($p<.20$) in Model 2. The final model included age, gender, race, randomization assignment, variables significant ($p<.20$) in prior models, and illness variables. All independent variables included in these models were those assessed at baseline except for the health insurance time-variable predictor. Because there could be disagreement among researchers as to whether a particular variable best belonged in a particular category, or with the

modeling strategy chosen, we entered all variables simultaneously and included them all in a multivariable model to assess the consistency of the findings.

Predisposing variables considered in these models included age, gender, race, marital status, birthplace, recent incarceration (five years), first language, family or friends using drugs, and family or friends supporting abstinence. Enabling variables included health insurance within six months prior to study enrollment, any visit to a medical clinic or private physician, or to an emergency department for medical care in the past six months, inability to get a regular doctor due to transportation problems, fear that others might find out about their health problems as a barrier to connecting with a regular doctor, the belief that the individual did not need a regular physician, alcohol and drug problems as measured by the total score of the InDUC-2L, injection drug use ever, current smoking, readiness to change (SOCRATES recognition and taking steps scales), and problem use of heroin, other opiates, and marijuana. Illness variables were physical health-related quality of life as measured by the SF-36 Physical Component Summary (PCS), the subject's belief that medical treatment is important, depressive symptoms (CES-D score), past suicide attempt or prescription of a medication for a psychiatric or emotional problem ever (from the ASI). Presence of any chronic medical illness replaced PCS as a dichotomous illness variable in a secondary analysis. To assess whether the association between these predictors and linkage varied by randomization group, we tested the interaction of each factor in the final model and randomization group in the HELP controlled trial, and repeated the multivariable model stratified by randomization group. Stratified results are presented only when relevant.

The Kaplan-Meier product limit estimator was used to estimate the unadjusted probability of linkage at a given time point (we chose 12 months, the midpoint of follow-up, to illustrate these proportions) for variables retained in the final model. Reported p-values are two-tailed, and a p-value of less than 0.05 was considered statistically significant. Analyses were carried out using *SAS/STAT* software (2001).

RESULTS

Of 470 subjects in the cohort, 2 died before follow-up and 400/468 (85 percent) completed at least one interview during the two-year follow-up period; there were 684 person-years of follow-up for the cohort. Subjects completed a mean of 17.5 months of follow-up (median 24 months). White

subjects were significantly less likely to be lost to follow-up (34 percent versus 54 percent), while subjects with family and friends that used drugs (71 percent versus 84 percent), and who believed medical treatment to be important (40 percent versus 58 percent), were less likely to be lost. Subjects lost to follow-up did not differ significantly by any other enabling, predisposing, or illness variable.

Subject Characteristics (Table 1)

Most of the 470 subjects were male, mean age was 36, and of the 11 percent for whom English was not their first language, 58 percent preferred to speak with their physician in English. Approximately 60 percent were employed and the same percentage had no health insurance in the past six months. Almost half were homeless. Most reported recent health care use.

Table 1: Characteristics of 470 Adults with No Primary Medical Care in Inpatient Detoxification

Characteristic	*Percent*
Male	76
Race/ethnicity	
White	46
Black	37
Hispanic	11
Other	6
U.S. born	87
English first language	89
Married	8
Unemployed (past six months)	39
Uninsured (past six months)	60
Homeless (> 1 night in the past six months)	47
Incarceration (past five years)	53
Recent health care use*	82
Recent medical visit**	26
Friends or family support abstinence	70
Friends or family use drugs	82
Current smoker	86
Problem alcohol use***	86
Problem heroin use***	38
Problem cocaine use***	75
Injection drug use (ever)	36
Chronic medical illness	47

*Addiction, mental health, episodic medical care, hospital, or emergency department visit in the past six months.

**At least one visit to a medical clinic or physician, not primary care, in the past six months.

***See methods for definition of problem use.

Many subjects reported barriers to linking with primary medical care. These barriers included inability to get to services due to transportation problems (28 percent), fear that others would find out about their health problems (11 percent), and not feeling that they need a regular physician (21 percent). On the other hand, 55 percent believed medical treatment was important.

Many subjects (69 percent) had problem marijuana use, and 87 percent had problem use of more than one substance (not including nicotine). The mean score on the recognition scale (possible range 7 to 35) of the SOCRATES was 33 +/−3 (SD) and on the taking-steps scale (possible range 8 to 40) was 36 +/−4 (SD).

With regard to medical and psychiatric illness, comorbidity was common. For example, 47 percent reported a chronic medical illness. The mean Short-Form Health Survey (SF-36) Physical Component Summary (PCS) score was 48+/−11 (SD) (50 is the mean score for the U.S. general population, which has an older mean age than study subjects) (Ware 1994). Depressive symptoms were very common (CES-D mean score 33+/−12); 90 percent had a CES-D score > 16, and 80 percent had CES-D score > 21, levels that correlate with a depression diagnosis (Radloff 1977). Other markers of psychiatric disease included the findings that 26 percent had ever been prescribed a medication for a psychiatric or emotional problem, and 22 percent had ever attempted suicide.

Linkage with Primary Medical Care

Of the 400 subjects with follow-up, 253 (63 percent) linked with primary medical care; 56 percent (109/195) of subjects in the nonenhanced (usual) primary care referral (control) group of the randomized HELP study linked with primary care. The annualized rate of linkage for the entire cohort was 53 linkage events per 100 person-years; the corresponding rate was 44 per 100 person-years in the control group of the HELP trial. The final multivariable model considering predisposing characteristics, enabling factors, and illness is reported in Table 2 along with unadjusted estimated predicted probabilities of linkage at one year, the midpoint of follow-up. All predictor variables were assessed during the baseline interview except for having health insurance during the follow-up period. Women, those with no recent incarceration, persons with support for abstinence by family or friends, and those who had visited a medical clinic or physician in the six months prior to study enrollment were significantly ($p < 0.05$) associated with shorter time to linkage

Table 2: The Relationship between Predisposing, Enabling, and Illness Variables and Linkage with Primary Medical Care over Two Years for Adults with Addictions ($n = 392$*)

	*Hazard Ratio (Adjusted) (95% Confidence Interval)***	*Predicted Probability of Linkage at Twelve Months[§] (Unadjusted for Other Factors)*	
Older age (years)	1.01 (0.99–1.03)	≥35 years	0.48
		< 35 years	0.41
Female sex	1.67 (1.23–2.28)	Female	0.52
		Male	0.42
Minority race/ethnicity	1.32 (0.98–1.78)	Minority	0.48
		White	0.39
Married	1.49 (0.95–2.36)	Married	0.51
		Not married	0.44
Incarceration (past five years)	0.73 (0.56–0.96)	Incarceration	0.38
		No incarceration	0.52
Family or friends support abstinence	1.60 (1.18–2.18)	Support	0.49
		Do not support	0.36
Medical visit (episodic, past six months)[++]	1.48 (1.10–1.97)	One or more	0.53
		None	0.42
Smoking (current)	1.41 (0.92–2.14)	Smoker	0.47
		Nonsmoker	0.30
Substance problem recognition[***]	1.04 (0.99–1.10)	≥34	0.44
		< 34	0.46
Belief that medical treatment is important	1.12 (0.81–1.56)	Important	0.46
		Not important	0.42
SF-36 PCS^	1.00 (0.98–1.01)	≥48.6	0.41
		< 48.6	0.49
CES-D score[+]	1.01 (0.99–1.02)	> 16	0.44
		≤16	0.47
Suicide attempt ever	0.95 (0.67–1.35)	Ever	0.49
		Never	0.44

continued

Table 2. *Continued*

	*Hazard Ratio (Adjusted) (95% Confidence Interval)***	*Predicted Probability of Linkage at Twelve Months[§] (Unadjusted for Other Factors)*	
Psychiatric medication ever	1.14 (0.81–1.60)	Ever prescribed	0.48
		Never prescribed	0.44
Insurance during follow-up	1.63 (1.19–2.22)	Insurance	0.54[††]
		No insurance	0.41
Randomization to enhanced primary care referral	1.78 (1.36–2.34)	Enhanced referral	0.52
		Usual care referral	0.37

All variables in the model were collected at the baseline research interview except insurance during follow-up, which was assessed at 6, 12, 18, and 24 months for each preceding 6-month period and entered as a time-variable predictor (see text).

Cut-point for continuous variables was the median value, except for the CES- D for which the clinically meaningful 16 or greater cutoff was used. See text for predicted probability of linkage by insurance during follow-up.

*Eight of four hundred subjects with follow-up in the study had missing data on one or more of the variables listed and are therefore excluded from these proportional hazards regression analyses.

**Adjusted for all other variables listed in the table.

***Recognition subscale of the SOCRATES (Stage of Change Readiness and Treatment Eagerness Scale).

^SF-36 PCS = Short Form Health Survey Physical Component Summary.

+CES-D = Center for Epidemiologic Studies Depression.

++At least one visit to a medical clinic or physician, not primary care, in the past six months.

§Based on the Kaplan-Meier product limit estimator.

††Predicted probability of linkage at six months associated with insurance during that six-month follow-up period.

with primary medical care in analyses adjusted for age, race/ethnicity, marital status, smoking, belief that treatment was important for their medical problem, physical health-related quality of life, depressive symptoms, past suicide attempt, and prescription of a medication for a psychiatric or emotional problem. In the same model, minorities, married persons, and those recognizing their addiction linked sooner to primary medical care at a borderline level of significance ($p<0.10$). In addition, although insurance at baseline was not significantly associated with linkage at follow-up (unadjusted Hazard Ratio 1.14, 95 percent CI 0.88–1.46), reported insurance status during follow-up was significantly associated with linkage to primary medical care during the same time period for which the insurance was reported (adjusted Hazard Ratio 1.63 [95 percent CI 1.19–2.22], where not having insurance is the reference group).

In a model in which SF-36 PCS score was replaced with the dichotomous indicator of chronic medical illness, results were similar. In a model that did not adjust for a recent episodic medical utilization, SF-36 PCS remained nonsignificant (adjusted HR 1.00, 95 percent CI 0.99–1.01). In a model that forced in these previously nonsignificant variables, alcoholism severity (adjusted HR 0.87, 95 percent CI 0.57–1.33), drug addiction severity (adjusted HR 0.50, 95 percent CI 0.19–1.34), addiction treatment (adjusted HR 1.10, 95 percent CI 0.82–1.48), or mental health visit (adjusted HR 0.94, 95 percent CI 0.62–1.45) in the past six months remained nonsignificant. Significant and nonsignificant variables in the final model using the sequential modeling approach remained so in a single model containing all of the independent variables.

There were no significant interactions between factors associated with linkage and randomized group in the HELP clinical trial except for a marginally significant interaction between randomization group and insurance during follow-up. In that multivariable model stratified by randomization group (e.g., enhanced referral intervention group versus nonenhanced [usual] primary care referral control group), the insurance effect was smaller among subjects in the intervention group (HR 1.30, 95 percent CI 0.85–1.98) than it was among subjects assigned to the control group (HR 2.45, 95 percent CI 1.44–4.16).

DISCUSSION

A substantial proportion of this relatively young cohort of addicted adults with high health care utilization but no existing regular primary medical care

relationship failed to link with primary medical care after residential detoxification. Women, those with recent episodic medical visits, family support for abstinence, and those with insurance after detoxification, were more likely to link with primary care. Recent incarceration decreased the likelihood of linkage.

Men with and without addictions are less likely to use primary medical care (Saitz, Mulvey, and Samet 1997; Lim et al. 2002; Gallagher et al. 1997). That men are less likely to link to care after detoxification suggests that interventions to improve linkage could target men when they are reachable in inpatient detoxification units. Many incarcerated adults report poor health status and failure to obtain needed medical care (Conklin, Lincoln, and Tuthill 2000). Since those with past incarceration were less likely to link with primary care after detoxification, efforts (already nascent in some communities [Conklin, Lincoln, and Flanigan 1998]) toward improving access to primary care should be studied.

Since our data suggest that prior contacts with episodic medical care enhance the likelihood of entering primary care after detoxification, these care sites could make linkage efforts standard practice. This finding is consistent with prior work finding that episodic medical illness is associated with having primary medical care in adults with addictions (Saitz, Mulvey, and Samet 1997). Our finding that social support for abstinence can increase linkage suggests that patients with little support could receive social support counseling, a method already known in other settings to improve follow-up ambulatory appointment-keeping (Tanner and Feldman 1998).

Health insurance during follow-up but not at the time of detoxification was one of the strongest predictors of linking with primary medical care. This was particularly true for subjects who had not received an enhanced referral to a primary medical care clinic that served patients regardless of ability to pay. In studies of other populations including those with addictions (Saitz, Mulvey, and Samet 1997; Bierman et al. 1999), having insurance is associated with use of medical services. But in this population of addicted adults, many people who had no primary medical care had health insurance (40 percent, Table 1). And having insurance at the time of contact with the detoxification unit was not enough to facilitate subsequent linkage with primary care. Only having health insurance at the right time—during the early period after detoxification when patients may begin to recognize and become concerned about medical symptoms as their sensorium clears and priorities change—was the predictor of importance.

Coincidentally, the period of follow-up in this study was a time when Massachusetts implemented a substantial Medicaid expansion (starting

July 1, 1997) (MassHealth 2002). And most of our subjects (90 percent) who had insurance in follow-up reported Medicaid as the insurer. Making health insurance coverage available to adults with addiction (a group disproportionately lacking primary medical care) at the right time (e.g., when they are more likely to access primary health care) is likely a generalizable strategy for improving receipt of primary care services (McCarthy et al. 2002). Differing findings regarding insurance at different times (at the time of detoxification and afterward) and during implementation of a statewide policy also demonstrate the importance of accounting for the dynamic nature of insurance coverage and changing policy in health services research.

We had hypothesized that ethnicity, recent addiction or mental health treatment utilization, addiction severity, health status, substance problem recognition, and perceived need for medical care would affect linkage, yet they did not. The association between minority race and linkage did not reach statistical significance but was in the same direction as has been previously reported for linkage with alcohol treatment (Kirchner et al. 2000). There were no discernible effects for mental health utilization or health status in our study. Health status was not associated with having a regular source of care in another study of a similarly vulnerable homeless population (Gallagher et al. 1997). This "need" or illness factor, generally associated with health care utilization (Bierman et al. 1999), and associated with having primary care for people with addictions (Saitz, Mulvey, and Samet 1997), may not have risen to the top of a priority list (Gallagher et al. 1997), or perhaps the need was met with episodic or emergency but not primary medical care. That patient beliefs about needing a physician did not lead to getting one, is also likely explained by a reordering of priorities (such as relapse or social, legal, or psychological needs) after leaving the residential detoxification facility. For addiction treatment utilization, addiction severity, and substance problem recognition, the effects were in the hypothesized direction but they did not reach statistical significance even in this large sample. In addition, the relatively low variability in the sample may explain why an expected association was not found (e.g., all had drug dependence severe enough to warrant inpatient detoxification).

The major strengths of this study were its focus on an understudied, reachable population in need, standardized prospective data collection with a high follow-up rate, and analyses based on theory. In addition, we used a broad definition of primary care based on how a physician functions in the eyes of the patient rather than based on how a health system categorizes them

(Starfield 1998); this deliberate choice makes it very likely that subjects reporting no primary medical care truly did not have it.

Limitations of our study include a 15 percent loss to follow-up that could have biased the results, however, the minimal losses and few differences in subject characteristics make this issue less of a concern. Assessment of primary medical care linkage by self-report may be a concern, but interview assessment of this outcome was a focus of the study; it was detailed, it referred to the recent past, and it was validated against administrative data. And recall for an event like a visit to a new primary care physician is more accurate than recall for less notable events (Means et al. 1989). Finally, the generalizability of our results may be limited to adults with addictions in similar low-socioeconomic-status detoxification and treatment programs typically found in cities in the United States.

Patients with addictions have many medical needs that go unaddressed (DeAlba, Samet, and Saitz in press; Saitz 2003). Regular primary medical care can improve their health care utilization and outcomes (Weisner et al. 2001). The challenge is to facilitate access to that care. Access could be improved by integrating primary care with addictions specialty care by having services onsite (Weisner et al. 2001), by better links between care sites (Samet, Friedmann, and Saitz 2001; Samet, Saitz, and Larson 1996), and by providing health insurance. Attention must also be paid to patient motivation and barriers to access (Teitelbaum et al. 1992).

Patients with addictions who have primary care physicians have already been described in the literature, and compared with those who do not (Saitz, Mulvey, and Samet 1997). But to our knowledge, no prospective study has reported on factors associated with linkage to primary medical care after detoxification for those without a physician. Detoxification is often not followed by addiction treatment or medical care (Samet et al. 2003; Mark et al. 2002). Thus detoxification presents an opportunity to reach patients without primary care, who could benefit from such care, and who may not seek it without facilitated access. Our data from this unique population do not simply mirror findings in the general population or even in persons with addictions (Saitz, Mulvey, and Samet 1997; Lim et al. 2002; Gallagher et al. 1997). In this setting, social support for abstinence, episodic medical care delivery contacts, incarceration, and insurance at the right time take on importance for patients with addictions who do not have, but who need primary medical care. Identification of these factors, and others significant in general and other vulnerable populations (e.g., gender), suggest clinical and policy interventions targeted to those at greatest risk as we have outlined in this discussion. We

anticipate that this knowledge of potentially modifiable factors that affect linkage with ongoing care could be used by health systems, detoxification or addiction treatment programs, and by researchers designing interventions to improve entry into primary care at a specific common point of contact with the health care system for patients with addictions.

REFERENCES

Andersen, R. M. 1995. "Revisiting the Behavioral Model and Access to Medical Care: Does It Matter?" *Journal of Health and Social Behavior* 36 (1): 1–10.

Andersen, R., and J. F. Newman. 1973. "Societal and Individual Determinants of Medical Care Utilization in the U.S." *Milbank Memorial Fund Quarterly: Health and Society* 51 (1): 95–124.

Bierman, A. S., T. A. Bubolz, E. S. Fisher, and J. H. Wasson. 1999. "How Well Does a Single Question about Health Predict the Financial Health of Medicare Managed Care Plans?" *Effective Clinical Practice* 2 (2): 56–62.

Cherry, D. C., C. W. Burt, and D. A. Woodwell. 2001. *National Ambulatory Medical Care Survey: 1999 Survey. Advance Data from Vital and Health Statistics, No 322.* Hyattsville, MD: National Center for Health Statistics.

Conklin, T. J., T. Lincoln, and T. P. Flanigan. 1998. "A Public Health Model to Connect Correctional Health Care with Communities." *American Journal of Public Health.* 88 (8): 1249–50.

Conklin, T. J., T. Lincoln, and R. W. Tuthill. 2000. "Self-reported Health and Prior Health Behaviors of Newly Admitted Correctional Inmates." *American Journal of Public Health* 90 (12): 1939–41.

DeAlba, I., J. H. Samet, and R. Saitz. In press. "Burden of Illness in Drug and Alcohol Dependent Persons without Primary Medical Care." *American Journal on Addictions.*

Diamant, A. L., R. H. Brook, A. Fink, and L. Gelberg. 2001. "Assessing use of Primary Health Care Services by Very Low Income Adults in a Managed Care Program." *Archives of Internal Medicine* 161 (9): 1222–7.

Folstein, M. F., S. E. Folstein, and P. R. McHugh. 1975. "'Mini-mental State': A Practical Method for Grading the Cognitive State of Patients for the Clinician." *Journal of Psychiatric Research* 12 (3): 189–98.

French, M. T., K. A. McGeary, D. D. Chitwood, and C. B. McCoy. 2000. "Chronic Illicit Drug Use, Health Services Utilization and the Cost of Medical Care." *Social Science and Medicine* 50 (12): 1703–13.

Friedmann, P., Z. Zhang, J. Hendrickson, M. Stein, and D. Gerstein. 2003. "Effects of Primary Medical Care on Addiction and Medical Severity in Substance Abuse Treatment Programs." *Journal of General Internal Medicine* 18 (1): 1–8.

Gallagher, T. C., R. M. Andersen, P. Koegel, and L. Gelberg. 1997. "Determinants of regular source of care among homeless adults in Los Angeles." *Medical Care* 35 (8): 814–30.

Gelberg, L., R. M. Andersen, and B. D. Leake. 2000. "The Behavioral Model for Vulnerable Populations: Application to Medical Care Use and Outcomes for Homeless People." *Health Services Research* 34 (6): 1273–302.

Hargraves, J. L., P. J. Cunningham, and R. G. Hughes. 2001. "Racial and Ethnic Differences in Access to Medical Care in Managed Care Plans." *Medical Care* 36 (5): 853–68.

Kirchner, J. E., B. M. Booth, R. R. Owen, A. E. Lancaster, and G. R. Smith. 2000. "Predictors of Patient Entry into Alcohol Treatment after Initial Diagnosis." *Journal of Behavior and Health Services Research* 27 (3): 339–46.

Laine, C., W. W. Hauck, M. N. Gourevitch, J. Rothman, A. Cohen, and B. J. Turner. 2001. "Regular Outpatient Medical and Drug Abuse Care and Subsequent Hospitalization of Persons Who Use Illicit Drugs." *Journal of the American Medical Association* 285 (18): 2355–62.

Levin, S. M., J. G. Trumble, M. Edmunds, J. M. Statman, and R. C. Petersen. 1993. "Perspectives on Linkage of Primary Health Care and Substance Abuse Treatment." *Journal of Addictive Diseases* 12 (2): 1–8.

Lim, Y. W., R. Andersen, B. Leake, W. Cunningham, and L. Gelberg. 2002. "How Accessible Is Medical Care in Homeless Women?" *Medical Care* 40 (6): 510–20.

Mark, T., J. Dilionardo, M. Chalk, and R. Coffey. 2002. *Substance Abuse Detoxification: Improvements Needed in Linkage to Treatment.* Report no. (SMA) 02-3728. Rockville, MD: Substance Abuse and Mental Health Service Administration.

MassHealth. 2002. "Demonstration Project Annual Report [accessed on April 13, 2002]." Available at http://www.state.ma.us/dma/researchers/res_pdf/1115_2001-demo4AR.pdf.

McCarthy, M. L., J. M. Hirshon, R. L. Ruggles, A. B. Docimo, M. Welinsky, and E. S. Bessman. 2002. "Referral of Mentally Uninsured Emergency Department Patients to Primary Care." *Academic Emergency Medicine* 9 (6): 639–42.

McGeary, K. A., M. T. French, S. Sacks, K. McKendrick, and G. De Leon. 2000. "Service Use and Cost by Mentally Ill Chemical Abusers: Differences by Retention in a Therapeutic Community." *Journal of Substance Abuse* 11 (3): 265–79.

McLellan, A. T., H. Kushner, D. Metzger, R. Peters, I. Smith, G. Grissom, H. Pettinati, and M. Argeriou. 1992. "The Fifth Edition of the Addiction Severity Index." *Journal of Substance Abuse Treatment* 9 (9): 199–213.

Means, B., A. Nigam, M. Zarrow, E. F. Loftus, and M. F. Donaldson. 1989. "Autobiographical Memory for Health-Related Events." *National Center for Health Statistics Vital Health Stat Series* 6 (2).

Miller, W. R., and J. S. Tonigan. 1995. *The Drinker Inventory of Consequences (DrInC): An Instrument for Assessing Adverse Consequences of Alcohol Abuse. Test Manual,* National Institute on Alcohol Abuse and Alcoholism Project MATCH Monograph Series, vol. 4. NIH publication no.95-3911. Bethesda, MD: National Institutes of Health.

———. 1996. "Assessing Drinkers' Motivation for Change: The Stages of Change, Readiness, and Treatment Eagerness Scale (SOCRATES)." *Psychology of Addictive Behavior* 10 (2): 81–9.

Morris, J. A. 1995. "Alcohol and Other Drug Dependency Treatment: A Proposal for Integration with Primary Care." *Alcoholism Treatment Quarterly* 13 (3): 45–56.

National Institute of Mental Health, the National Institute on Drug Abuse, and the National Institute on Alcohol Abuse and Alcoholism. 2003. Conference on "Expanding Mental Health, Drug and Alcohol Services Research Outside the Specialty Care System [accessed August 8, 2003]," March 10–13. Available at http://www.nimh.nih.gov/events/mhsrconf2003.cfm. Washington, DC: NIMH, NIDA, and NIAAA.

O'Connor, P. G., A. H. Oliveto, J. M. Shi, E. G. Triffleman, K. M. Carroll, T. R. Kosten, B. J. Rounsaville, J. A. Pakes, and R. S. Schottenfeld. 1998. "A Randomized Trial of Buprenorphine Maintenance for Heroin Dependence in a Primary Care Clinic for Substance Users versus a Methadone Clinic." *American Journal of Medicine* 105 (2): 100–5.

O'Malley, S. S., B. J. Rounsaville, C. Farren, K. Namkoong, R. Wu, J. Robinson, and P. G. O'Connor. 2003. "Initial and Maintenance Naltrexone Treatment for Alcohol Dependence Using Primary Care versus Specialty Care. A Nested Sequence of Three Randomized Trials." *Archives of Internal Medicine* 163 (14): 1695–704.

Radloff, L. S. 1977. "The CES-D Scale: A Self-report Depression Scale for Research in the General Population." *Applied Psychological Measurement* 1 (3): 385–401.

Saitz, R. 2003. "Medical Issues in Persons with Addictions." In *Principles of Addiction Medicine*, pp. 1027–52. Chevy Chase, MD: American Society of Addiction Medicine.

Saitz, R., K. P. Mulvey, and J. H. Samet. 1997. "The Substance Abusing Patient and Primary Care: Linkage via the Addiction Treatment System?" *Substance Abuse* 18 (4): 187–95.

Samet, J. H., P. D. Friedmann, and R. Saitz. 2001. "Benefits of Linking Primary Medical Care and Substance Abuse Services: Patient, Provider, and Societal Perspectives." *Archives of Internal Medicine* 161 (1): 85–91.

Samet, J. H., M. J. Larson, N. J. Horton, K. Doyle, M. Winter, and R. Saitz. 2003. "Linking Alcohol and Drug Dependent Adults to Primary Medical Care: A Randomized Controlled Trial of a Multidisciplinary Health Intervention in a Detoxification Unit (The Health Evaluation and Linkage to Primary care [HELP] Study)." *Addiction* 98 (4): 509–16.

Samet, J. H., R. Saitz, and M. J. Larson. 1996. "A Case for Enhanced Linkage of Substance Abusers to Primary Medical Care." *Substance Abuse* 17 (4): 181–92.

SAS Institute. 2001. *SAS/STAT*, statistical software (version 8.2). Cary, NC: SAS Institute.

Schlenger, W. E., L. A. Kroutil, and E. J. Roland. 1992. *National Evaluation of Models for Linking Drug Abuse Treatment and Primary Care: Linking Drug Treatment and Primary Care: A Casebook*. Report prepared for the National Institute on Drug

Abuse under contract no. 283-90-0001. Research Triangle Park, NC: Research Triangle Institute.

Starfield, B. 1998. *Primary Care: Balancing Health Needs, Services and Technology.* New York: Oxford University Press.

Stein, M. D., K. P. Mulvey, A. Plough, and J. H. Samet. 1998. "The Functioning and Well Being of Persons Who Seek Treatment for Alcohol and Drug Problems." *Journal of Substance Abuse* 10 (1): 75–84.

Tanner, E. K., and R. H. L. Feldman. 1998. "Strategies for Enhancing Appointment-Keeping in Low-Income Chronically Ill Clients." *Nursing Research* 46 (6): 342–4.

Teitelbaum, M., A. Walker, M. Gabay, M. Gross, T. Mason, and E. Weitzman. 1992. *Analysis of Barriers to the Delivery of Integrated Primary Care Services and Substance Abuse Treatment: Case Studies of Nine Linkage Program Projects.* Rockville, MD: Health Resources and Services Administration and ABT Associates.

Volpicelli, J. R., A. I. Alterman, M. Hayashida, and C. P. O'Brien. 1992. "Naltrexone in the Treatment of Alcohol Dependence." *Archives of General Psychiatry* 49 (11): 876–80.

Ware, J. E. Jr. 1993. *SF-36 Health Survey. Manual and Interpretation Guide.* Boston: The Health Institute.

———. 1994. *SF-36 Physical and Mental Health Summary Scales: A User's Manual.* Boston: The Health Institute.

Weisner, C., J. Mertens, S. Parthasarathy, C. Moore, and Y. Lu. 2001. "Integrating Primary Medical Care with Addiction Treatment." *Journal of the American Medical Association* 286 (14): 1715–23.

Willenbring, M. L., and D. H. Olson. 1999. "A Randomized Trial of Integrated Outpatient Treatment for Medically Ill Alcoholic Men." *Archives of Internal Medicine* 159 (16): 1946–52.

30 DRUG ABUSE AND DEPENDENCE

Jeffrey H. Samet

Drug abuse increasingly is recognized as an important mainstream health problem as a consequence of several factors: Injection drug use remains a major transmission risk for human immunodeficiency virus (HIV) infection; more than 1 million drug arrests occur in the United States each year; and costs are enormous, estimated as greater than $110 billion in the United States in 1995. About 15 million people older than 12 years of age have used illicit drugs at least once during the past month, and about 3.5 million people are classified as drug dependent. Medical complications of drug abuse are predominantly infectious but span organ systems and range from cocaine-related cardiac arrhythmia to neuropsychiatric effects of hallucinogens.

Definitions

The terms *drug* (or substance) *dependence* and *drug abuse* have specific clinical meanings (Table 30–1). Dependence is the more severe disorder and frequently is associated with physiologic and psychological manifestations. *Tolerance* and *withdrawal* are the major physiologic manifestations of drug dependence. Tolerance is defined as either a need for increased amounts of the substance to achieve the desired effect or a diminished effect with continued use of the same amount of the substance. Withdrawal is manifested by a characteristic syndrome with sudden abstinence, but it may be relieved or avoided if the same or a closely related substance is taken. The other criteria for dependence relate to the pattern of drug use (i.e., taken in a larger amount or longer period than intended); effects on life activities (i.e., great deal of time spent on activities to obtain, use, or recover from the substance; reduction in social, occupational, or recreational activities as a result of substance use); and the psychological need to use the substance (i.e., use despite awareness of adverse consequences, persistent desire for the substance, or inability to control its use).

A diagnosis of substance abuse requires the recurrent use of a substance over 12 months with subsequent adverse consequences (e.g., failure to fulfill a major role at work, school, or home; legal problems; persistent interpersonal problems) or placement of an individual in high-risk, physically hazardous situations. Addiction is a chronic, relapsing illness characterized by compulsive drug seeking and use.

The degree of harm associated with occasional drug use or "experimentation" is difficult to quantify, and no definition has been assigned formally to the use of illicit drugs with consequences less than those associated with the abuse definition. Fear of progression to abuse or dependence, the potential morbidity of use of drugs such as cocaine, the criminality associated with drug use, and the high-risk behavior while under the influence of a drug are the basis of recommendations to proscribe use of these substances.

Table 30–1 • DIAGNOSTIC CRITERIA FOR DEPENDENCE AND DRUG ABUSE

DEPENDENCE (≥3 NEEDED)	ABUSE (≥1 FOR 12 MO)
1. Tolerance	1. Recurrent substance use resulting in failure to fulfill major role obligations at work, school, or home
2. Withdrawal	2. Recurrent substance use in situations in which it is physically hazardous
3. Substance often taken in greater amounts or over longer period than intended	3. Recurrent substance-related legal problems
4. Persistent desire or unsuccessful efforts to cut down or control substance use	4. Continued substance use despite having persistent or recurrent social or interpersonal problems caused or exacerbated by the effects of the substance
5. Great deal of time spent in activities necessary to obtain the substance, use the substance, or recover from its effects	5. Never met criteria for dependence
6. Important social, occupational, or recreational activities given up or reduced because of substance use	
7. Continued substance use despite knowledge of having had persistent or recurrent physical or psychological problems that are likely to be caused or exacerbated by the substance	

Etiology

A minority of people who ever experiment with an illicit drug progress to a clinical drug abuse diagnosis. The cofactors responsible for progression to dependence and abuse are only partially defined. Genetic susceptibility, social context of the drug use, and comorbid psychiatric conditions are considered important factors affecting an individual's potential for subsequent problems. Twin studies suggest that genetics plays a role in a person's positive or negative perception of a drug's effect. The social context in which drug abuse develops and is expressed is important. Returning Vietnam War veterans addicted to heroin were relatively easy to treat compared with addicts on the streets of the United States, in part because the veterans had become addicted in a setting different from the one they found on return home and were exposed to few enduring environmental cues. Psychiatric comorbidity, particularly depression and panic disorders, seems to be a high-risk condition for the development of drug abuse and possible consequences of this abuse.

DRUG OF ABUSE: HEROIN AND OTHER OPIOIDS

Classification

Opioids, including naturally occurring alkaloids (opiates derived from the poppy plant *Papaver somniferum*), semisynthetic compounds (chemically altered alkaloids), and synthetic agents, are potent analgesics and produce an intense euphoria associated with nausea; drowsiness; miosis; and a decrease in respiration, pulse, and blood pressure. Opioids also are valued for their calming, antitussive, and antidiarrheal properties. Depending on the particular effect on opioid cell membrane receptors, they may be classified as agonists (morphine, heroin, methadone), partial agonists-antagonists (buprenorphine), or antagonists (naloxone, naltrexone). These drugs have led to many medical complications because of their abuse potential and their parenteral route of administration.

History

In the 19th century, opioids were used commonly in many settings in the United States. The drug was supplied freely by physicians to treat symptoms of pain, anxiety, cough, and diarrhea. Opiates also were available without restriction in commercial medicinal remedies.

In 1806, a pure substance was isolated from opium and named *morphine* after the Greek god of dreams Morpheus. By the middle of the 19th century, the advent of the hypodermic needle allowed this inexpensive, standard-strength agent to become a highly effective pain-killing and calming therapy. Smoking opium, which has no medicinal value, also rose in the latter half of the 19th century. In 1898, heroin was introduced commercially by the Bayer Company as an antitussive and was used as therapy for morphine addiction. The increasing recognition of the perils of opiate addiction, its identification with foreign groups and internal minorities, and concern over the estimated prevalence of 250,000 opiate users in 1900 led to a series of state and federal measures culminating in the Harrison Narcotic Act in 1914, which legislated controls over the importation and distribution of opiates.

Opiate use remained a problem in the early 20th century despite interdiction efforts and the development and dismantling of narcotic clinics that maintained narcotic addicts with prescription drugs. In the 1920s, narcotic abuse became a predominantly underground activity. Efforts to treat narcotic addiction as a medical problem were limited until the advent of methadone maintenance therapy in the 1960s.

Epidemiology

In the United States, an estimated 3 million people have reported prior use of heroin. About 170,000 of the estimated 810,000 opioid-dependent persons are enrolled in opioid treatment programs.

An estimated 150,000 individuals become new heroin users each year, an upward trend comparable with increases seen in the epidemic associated with the Vietnam War in the late 1960s. New users are likely to be young (72% <26 years old), noninjecting (63%), and urban dwellers (89%). Polysubstance abuse is increasingly common, with

50% of male and 25% of female narcotic addicts meeting the criteria for alcohol dependence. Nicotine is the most common substance used together with opiates.

Biomolecular Mechanisms of Action

Opioids exert their effects on specific receptors for three distinct families of endogenous opioid peptides: enkephalins, endorphins, and dynorphins. In the central nervous system, three major classes of opioid receptors with unique selectivity and pharmacologic profiles have been identified: μ, κ, and δ. Subtypes of these major classes (μ1, μ2, κ1, κ2, κ3, δ1, δ2) have been elucidated primarily by the use of selective receptor antagonists. μ receptor activity is associated with the most prominent manifestations of morphine and heroin: respiratory depression, analgesia, euphoria, and the development of dependence. It is thought that opioid peptides acting as neurotransmitters or neuromodulators exert their actions at neuronal synapses.

Clinical Pharmacology

Heroin may be injected intravenously or subcutaneously, snorted, smoked, or ingested. The parenteral and inhaled routes of administration result in the most rapid delivery of drug to the brain and are the most potentially addicting. As the purity of street heroin has increased from less than 5% in the 1960s to 80% in the 1990s, its nonparenteral administration has risen. Heroin may be used intermittently or regularly. Intermittent users generally either quit or become regular users within 1 to 3 years. Given the drug's short half-life, regular users require two to four daily doses to avoid withdrawal symptoms.

Heroin's initial effect is an intense euphoria described as a "rush" or "kick," compared in intensity and pleasure with an orgasm, that lasts 45 seconds to several minutes. The initial effects may be perceived as a turning in the stomach with tingling and warmth. A user's first experience may be unpleasant because of nausea, vomiting, and anxiety, but these effects decrease or become less of a concern to the user over time. The intense euphoria is followed by an intoxicated pleasant feeling referred to as "nodding," with decreased respiration and peristalsis. The depressant effect of heroin on the central nervous system is marked, particularly after parenteral administration. Sedation, mental clouding, decreased visual acuity, heavy feeling in the extremities, light sleep with vivid dreams, and reduction in anxiety are typical, at least until tolerance develops. Physical signs include miosis, decreased heart rate, and lowered blood pressure. In addition to these effects on opioid receptors, heroin causes the release of histamine, which may result in itching, scleral injection, and hypotension.

Table 30–2 • SIGNS AND SYMPTOMS OF WITHDRAWAL FROM OPIOIDS AND COCAINE

OPIOID WITHDRAWAL	
Vital signs	Tachycardia, hypertension, fever
Central nervous system	Craving, restlessness, insomnia, muscle cramps, yawning, miosis
Eyes, nose	Lacrimation, rhinorrhea
Skin	Perspiration, piloerection
Gastrointestinal	Nausea, vomiting, diarrhea
COCAINE WITHDRAWAL	
Crash	Depression, fatigue
Withdrawal	Anxiety, high craving
Extinction	Normalization of mood, episodic craving

High levels of tolerance develop rapidly with regard to respiratory depression, analgesia, sedation, vomiting, and euphoric properties. Little tolerance develops for miosis or constipation, so a heroin addict with an acutely painful medical condition may complain of insufficient analgesia despite pinpoint pupils. Cross-tolerance is common among opioids.

From the patient's perspective, withdrawal from heroin is a dreaded clinical condition, a mix of emotional, behavioral, and physical signs and symptoms (Table 30–2). Although unpleasant, it is not life-threatening. The timing of withdrawal symptoms, which are related directly to clearance of the drug, begins 4 to 8 hours after the last dose of heroin. The acute withdrawal syndrome peaks in intensity after 36 to 72 hours and resolves over 5 to 7 days.

In addition to the acute abstinence syndrome, a protracted abstinence syndrome occurs and lasts 6 months or more. In contrast to the hyperadrenergic characteristics of the primary withdrawal syndrome (tachycardia, hypertension, elevated temperature, mydriasis, and diaphoresis), the period afterward can consist of sluggishness, sleep disturbance, and malaise. Craving can recur for years after cessation of drug use. An understanding of the nature of recovery from heroin use is important for setting appropriate expectations for the patient and the health care provider.

Clinical Complications

Deaths from heroin overdose have been increasing in several countries. Nevertheless, most opioid-related medical complications occur as a result of the spread of infectious agents by injection drug use among heroin addicts (Fig. 30–1). The manifestations of these medical complications are protean but frequently nonspecific, such as fever, malaise, weight loss, pain, and dyspnea. The underlying causes include endocarditis, cellulitis, HIV disease, hepatitis, pneumonia, and a variety of abscesses. Additionally, specific syndromes have been attributed to direct toxic effects of the opioid itself.

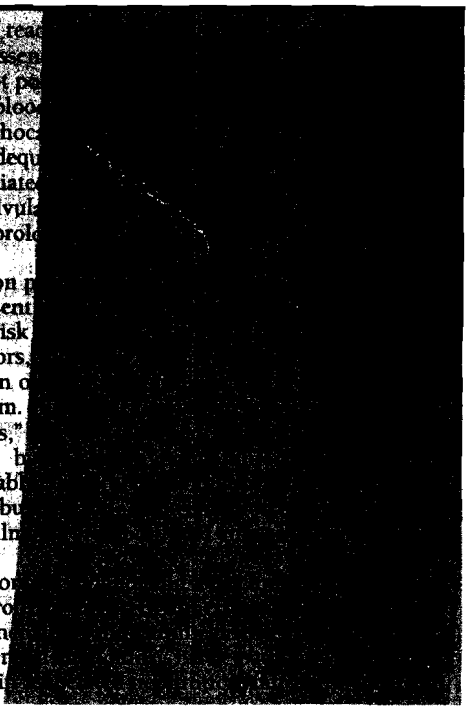
Reports of infectious complications of intravenous drug use before 1970 included falciparum malaria, tetanus, endocarditis, and acute hepatitis. Although falciparum malaria and tetanus are now rare, hepatitis B and C are exceedingly common and serologically detectable in most heroin addicts in the United States and abroad.

The major cardiac complication of opioid abuse is bacterial endocarditis (Chapter 310) caused by injection drug use. *Staphylococcus aureus* is the most frequently reported bacterial isolate, and the tricuspid valve is the most common valve involved. Left-sided valvular infection is associated with a worse prognosis, as are the uncommon gram-negative and fungal infections.

Opioid abusers normally have acute rather than subacute endocarditis. The initial clinical finding can be fever alone in half the cases, or fever may be associated with pulmonary infiltrates from right-sided emboli or systemic embolic phenomena, such as arthritis, abscess, and osteomyelitis. The diagnosis of endocarditis in a febrile injection drug user is difficult because of the poor sensitivity and specificity of rea
Blood cultures are esse
tient follow-up is not pe
mended until initial bloo
and specificity of echoc
studies but are not adequ
complications associate
diomyopathy, perivalvul
system such as QT prol
pulmonale.

The most common p
monia, which is present
ated for fever. The risk
combination of factors,
function, suppression o
of clouded sensorium.
"talc granulomatosis,"
granulomas caused b
substances, most notabl
ciated with opioid abu
chospasm, septic puln
mediastinitis.

Renal complicatio
(myoglobinuria, necro
with endocarditis or h
drome, renal failure, r
monly found in heroi



Continued

diffuse glomerulosclerosis (Chapter 119). In HIV-infected patients, HIV-associated nephropathy also is found (Chapter 420).

Of patients in methadone maintenance clinics, 50 to 90% have positive serologic studies for hepatitis B and C. Complications of these infections (Chapter 152) range from chronic asymptomatic antigenemia to cirrhosis and hepatocellular carcinoma.

Neurologic complications of opioid abuse are infectious and noninfectious. Seizures, most often generalized, are the most common noninfectious complication. The cause of seizures includes overdose, with centrally mediated respiratory depression and hypoxia, and cerebral infarction. Meningitis, mycotic aneurysm, and abscesses (epidural, subdural, and brain) are well-described infectious conditions resulting from injection drug use. In HIV-infected patients, HIV-associated neurologic infectious and noninfectious diseases occur (Chapter 414).

Psychiatric conditions among opioid abusers are common and include alcohol abuse/dependence, major depression, phobic disorders, and antisocial personality, all of which have a greater than 15% lifetime prevalence. Men are four to seven times more likely to have an antisocial personality than women are; women more commonly have depression. Women abusers are at high risk of being victims of violence.

Immunologic abnormalities among heroin addicts are described before the acquired immunodeficiency syndrome epidemic. Chronic morphine decreases the [illegible] an opiate antagonist, can [illegible] globulinemia of addicts, [illegible] genic stimulation, is th[illegible] false-positive indirect sy[illegible] consequences of opioid-[illegible] The most prominent cli[illegible]

The associated medi[illegible] users mirror those of no[illegible] with a few caveats. HIV[illegible] quency of bacterial pneu[illegible] sarcoma. HIV testing w[illegible] ommended strongly for[illegible] promote the use of cond[illegible] can reduce HIV transm[illegible] cacious in reducing th[illegible]

FIGURE 30–1 • Injection drug use typically leads to this appearance, which results from repeated superficial thrombophlebitis of accessible veins in the arm or elsewhere in the body. Sharing and reuse of syringes and needles puts these patients at risk of a wide range of infections, including bacterial septicemia, systemic fungal infection, hepatitis B, hepatitis C, and HIV infection. Right-sided endocarditis is a common complication. (From Forbes CD, Jackson WF: Color Atlas and Text of Clinical Medicine, 3rd ed. London, Mosby, 2003, with permission.)

PRESCRIPTION OPIOIDS

Opioids are prescribed appropriately for acute and chronic pain. These analgesic medications vary in their potency and bioavailability and include oxycodone (Percodan, Percocet), hydrocodone (Vicodin), hydromorphone (Dilaudid), and meperidine (Demerol). A new slow-release formulation of oxycodone, OxyContin, has gained great popularity among recreational and dependent opioid users since its approval in 1995. By crushing these pills and destroying the slow-release matrix in which the opioid is embedded, oxycodone is abused orally, parenterally, and intranasally, yielding effects comparable to heroin. An estimated 2.6 million Americans misuse pain relievers each year. Despite these abuses, it is important for the physician to understand that it is uncommon for appropriate use of these medications in the treatment of pain to lead to opioid dependence (Chapter 29).

DRUG OF ABUSE: COCAINE AND OTHER PSYCHOSTIMULANTS

Classification

Cocaine, an alkaloid extracted from coca leaves, and other psychostimulants (e.g., amphetamine, methamphetamine) rapidly increase the concentration of several neurotransmitters in synaptic junctions and stimulate the sympathetic and central nervous systems. Topical cocaine is used in otolaryngologic procedures, and psychostimulants are used either for their stimulant effects or for their paradoxical calming effect in some patients with attention-deficit hyperactivity disorder.

History

The earliest recorded use of cocaine in the form of ingested coca leaf occurred around 3000 BC. In 1860, cocaine was isolated and incorporated into tonics, teas, and wines. In the 1880s, an Atlanta druggist patented a product that contained two naturally occurring stimulants, cocaine and caffeine, which eventually became known as Coca-Cola; until 1903, it contained approximately 60 mg of cocaine per 8-oz serving. In the late 19th century, reports of cocaine addiction surfaced, and its use was restricted after passage of the Harrison Narcotic Act of 1914. The abuse potential of amphetamines led to their being listed as schedule II drugs, which are defined as having a high potential for abuse with severe liability to cause psychic or physical dependence.

Epidemiology

An estimated 1.5 million Americans, representing 0.8% of the population aged 12 years and older, have used cocaine in the past month. More than 900,000 Americans use cocaine for the first time each year, and more than 30 million Americans have used cocaine at least once. Use is higher in the 18- to 34-year-old age group (1.5 to 2.0%), in men than in women (1.1 versus 0.5%), in urban areas, and among individuals with less education. Although current cocaine use is highest in the unemployed (2.4%), 73% of adult users are employed full-time or part-time. Current cocaine use is similar for whites (0.8%), blacks (1.0%), and Hispanics (1.1%).

Biomolecular Mechanisms Of Action

Cocaine increases neurotransmitter concentrations at the synaptic terminal by blocking the reuptake of norepinephrine, dopamine, and serotonin and by potentiating the release of these monoamines. In the heart, α-adrenergic and β-adrenergic receptors are stimulated.

past year, and more than 75 million have used marijuana in their lifetime.

Cannabinoids bind to specific receptors for the endogenous ligand anandamide: CB1 in the brain and CB2 in the periphery. G-protein activation occurs as a result of the receptor binding and has three effects: inhibition of adenylate cyclase, increased potassium ion conductance, and decreased calcium ion conductance. CB1 receptors are concentrated in the globus pallidus, hippocampus, cerebral cortex, cerebellum, and striatum.

Smoked marijuana results in a variety of acute changes within 3 minutes that peak within 20 to 30 minutes; when ingested, onset takes 30 to 60 minutes, and the peak effect occurs after 2 to 3 hours. An average cigarette contains 2.5 to 5 mg of THC, and 50 to 60% of it is absorbed. THC is lipophilic and distributed rapidly throughout the body. Because of slow release from adipose tissue, THC or its metabolites can be found in urine 1 to 3 days after use in nonchronic users and 30 days after use in chronic users (see Table 30–3).

Most effects last 2 to 3 hours after inhalation; psychomotor effects can last 11 hours. Effects include conjunctival injection, mild euphoria, impaired memory, dry mouth, motor incoordination, time-space distortion, increased visual and auditory awareness, increased hunger, sleepiness, and spontaneous laughter; some may experience nausea, headaches, tremors, decreased muscle strength, and increased anxiety. Few chronic effects have been attributed to marijuana use, but an amotivational syndrome has been described in which young people lose goal-directed behavior with regard to school or work.

DRUG OF ABUSE: LSD AND OTHER HALLUCINOGENS

Hallucinogen use results predominantly in changes in thought, perception, and mood. Minimal impairment occurs in memory or intellect. This class of drugs is not generally associated with stupor, narcosis, or excessive stimulation. Users do not exhibit craving. The two major categories of hallucinogens are indolamines (e.g., lysergic acid diethylamide [LSD], dimethyltryptamide, psilocybin) and phenylethylamines (e.g., methylenedioxyamphetamine, methylenedioxymethamphetamine [MDMA], mescaline). Related drugs include phencyclidine (PCP), nutmeg, morning glory seeds, catnip, nitrous oxide, and amyl or butyl nitrite. These drugs have no appropriate clinical role.

In the United States, the lifetime prevalence of hallucinogen use is about 11%; more than 25 million individuals have used hallucinogens at least once. LSD was used widely on college campuses in the 1960s. The 1970s and early 1980s saw a decline in the use of most hallucinogens. Hallucinogen use increased in the 1990s, however, and by 1998 there were 1.2 million new users, twice the average in the 1980s.

The classic hallucinogens are structurally similar to many major neurotransmitters, but serotonin (5-hydroxytryptamine [5-HT]) agonist or partial agonist properties have been associated most consistently with its actions. These drugs bind at 5-HT_{2A} and 5-HT_{2C} receptors with high affinity. These receptors are found in greatest density in brain cortical regions (cerebral cortex, claustrum, caudate putamen, globus pallidus, ventral pallidum, islands of Calleja, mammillary nuclei, and inferior olive) and may have a role in depression and suicide.

Hallucinogen use results in an altered perception of one's environment marked by a subjective feeling of enhanced mental activity, perceptual distortions, visual hallucinations, sharpened sense of hearing, and a reduced ability to tell the difference between one's self and one's surroundings. These drugs can produce sympathomimetic effects, including mydriasis, flushed face, fine tremor, piloerection, high blood pressure, hyperthermia, and hyperglycemia. Panic attacks and psychosis are the two major adverse effects. Clinically "desired" effects and adverse effects vary by specific hallucinogen. Altered perceptions can be associated with paranoid delusions, manic or depressed behavior, and confusion. Aggressive behavior has been described with psychosis; in particular, PCP has been implicated in violent crimes. The psychotic episodes can last hours or days, and flashbacks can occur. Precipitants for flashbacks are anxiety, stress, fatigue, emergence into a dark environment, and marijuana.

Although tolerance can develop with hallucinogens, the clinical syndrome is unusual inasmuch as chronic use is uncommon. No clinically significant withdrawal symptoms are known. Concerns about chronic use include prolonged psychotic episodes, decreased intellect, organic brain syndrome, and possibly "chromosomal damage," although definitive correlations have not been established.

The use of hallucinogens may be detected in the acute setting when examining a patient with toxic manifestations or may be noted when obtaining a history of drug use. After diagnosis, it is important to obtain a history of other substance abuse and psychiatric illness and a neurologic evaluation. No specific laboratory tests are required; a urine toxicologic screen for other drugs of abuse is recommended (see Table 30–3).

LSD often is sold as postage stamp–size papers impregnated with varying doses of LSD, from 50 μg to more than 300 μg. Doses of 20 μg can lead to psychological effects, with doses of 100 μg causing hallucinogenic psychoactive manifestations within 1 to 2 hours. Clearing of symptoms begins in 10 to 12 hours, although symptoms of fatigue and tension can persist for an additional 24 hours.

PCP, also known as "phencyclidine" or "angel dust," originally was developed as an anesthetic in the 1950s but was abandoned because of frequent postoperative delirium and hallucinations. It can be obtained in various forms (powder, liquid, tablet, capsule, or sprayed on other drugs such as marijuana) and administered by several routes (smoked, ingested, snorted, or injected intravenously). The drug is water soluble and lipophilic, so it penetrates fat stores and has a long half-life, up to 3 days. Casual use by smoking on a weekly basis is most common, although some have reported continuous intake lasting 2 days or longer. A pronounced pharmacologic characteristic of PCP is its analgesia and amphetamine-like stimulation in addition to hallucination. Ataxia, slurred speech, nystagmus, and numbness commonly are observed at doses of 1 to 10 mg. Emotional withdrawal, catatonic posturing resembling schizophrenia, and physical violence can result from its use.

DRUG OF ABUSE: BENZODIAZEPINES AND OTHER SEDATIVES

Benzodiazepines and the less commonly used barbiturates are legitimate therapeutic drugs with abuse potential. These drugs are designated as schedule IV substances by the Drug Enforcement Agency and the Food and Drug Administration. Schedule IV drugs have a low potential for abuse and lead to limited physical or psychological dependence.

Nonmedical use of tranquilizers and sedatives occurs in fewer than 2% of U.S. adults annually. The magnitude of the problem is substantially less than that of opioids, psychostimulants, and marijuana. This problem occurs largely in individuals who also abuse other substances. This finding is consistent with the experience in laboratory animals, which do not exhibit repeated self-administration, a standard measure of addictive potential, when exposed to benzodiazepines.

All benzodiazepines studied are capable of producing physiologic dependence even when used in low doses over prolonged periods as may be seen in clinical practice. The key to the diagnosis of benzodiazepine or other sedative abuse is evidence of inappropriate drug-taking behavior, including escalation in dose, obtaining prescriptions from multiple physicians, or taking the drug for reasons other than those for which it was prescribed. Physiologic dependence should not imply that inappropriate drug-taking behavior exists. Before initiating clinical use of benzodiazepines and other sedatives, a careful medical history must be obtained regarding current and prior substance abuse. Although not absolutely contraindicated, particular caution and extra monitoring are appropriate in patients with such a history.

NEW DRUGS OF ABUSE: CLUB DRUGS

Newer drugs of abuse, such as 3,4 methylenedioxymethamphetamine (MDMA), γ-hydroxybutyrate (GHB), and ketamine, are used in a variety of settings. When ingested in association with inadequate fluid intake, vigorous exercise, or a hot, humid environment, these drugs are particularly likely to cause complications.

MDMA (ECSTASY)

MDMA, commonly referred to as "Ecstasy," is a synthetic analogue of amphetamine and shares properties with amphetamine and hallucinogenic drugs. It acts on the serotonin transporter, stimulating serotonin release and inhibiting its reuptake. Although usually taken in

the pill form, MDMA also can be snorted, injected, or administered per rectum. The purity of MDMA tablets may vary 70-fold, and tablets may include caffeine, heroin, or mescaline. More than 8% of high school seniors report using MDMA.

MDMA's clinical effects are predominately sensory enhancement with distortion and decreased inhibitions. The onset of action is 30 to 60 minutes, the peak effects occur at 90 minutes, and the duration is 8 hours or more. Common adverse effects, which are similar to effects found with amphetamines and cocaine, include sweating, muscle spasms, involuntary teeth clenching, faintness, chills, and tachycardia. Psychological manifestations include confusion, depression, sleep problems, severe anxiety, and paranoia. High temperatures and muscle exertion from dancing seem to lower the threshold for serious MDMA-associated adverse effects, especially rhabdomyolysis; other reported adverse effects in the club setting include hyponatremia, dehydration, hypothermia, hypertensive crisis, and cardiac arrhythmias.

GHB

GHB, or "liquid ecstasy," is a metabolite of the neurotransmitter γ-aminobutyric acid. It is thought to function as a neurotransmitter, producing a dopaminergic response and release of an opiate-like substance. Its half-life is approximately 30 minutes. GHB is used for its euphoric and anabolic effects. Behavioral changes include increased aggression, and neurologic changes range from mild ataxia to apnea. Withdrawal symptoms are similar to those of sedative abuse and persist for 3 to 7 days. There is no antidote for GHB overdoses, and treatment is limited to nonspecific supportive care.

KETAMINE

Ketamine, commonly referred to as "Special K," is a fast-acting intravenous or intramuscular anesthetic that delivers hypnotic, analgesic, and amnesic effects. Most of ketamine's activity is associated with *N*-methyl-D-aspartate receptors. Because it causes an intense dissociative state and loss of physical control, ketamine use is associated with a high risk of injuries. Ketamine usually is acquired from veterinary clinics.

TREATMENT OF DRUG ABUSE AND PREVENTION OF RELAPSE

Patients who use illicit drugs benefit from treatment if they recognize that their substance use is a problem. The transtheoretical model considers a patient on a continuum from precontemplation (denial) toward maintenance (abstinence/recovery) (Fig. 30–2). The clinical approach should be tailored to the patient's readiness to change behavior and enter treatment. For all abused drugs, medical follow-up after any acute toxic presentation is essential to address substance abuse issues and possible coexisting medical and psychiatric problems.

The major goals of drug abuse treatment are detoxification, abstinence initiation, and relapse prevention. Treatment can be pharmacologic and nonpharmacologic. Pharmacologic approaches are offered by physicians specializing in addiction and increasingly by primary care physicians.

Some form of psychosocial treatment is the backbone of substance abuse treatment, be it psychotherapy, behavioral therapy, or counseling. Issues addressed in these encounters include teaching coping skills, changing reinforcement contingencies, fostering management

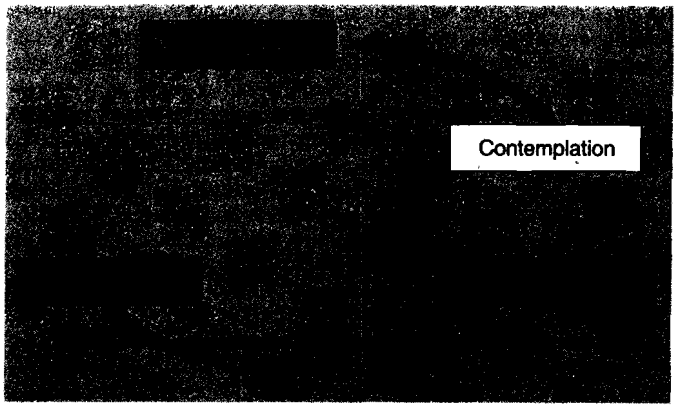


FIGURE 30–2 • Graphic depiction of the Prochaska and DiClemente model for readiness for behavioral change.

of painful effects, addressing motivation, improving interpersonal functioning, enhancing social supports, and encouraging compliance with and retention in pharmacotherapy. Much of this work is done by substance abuse care providers. Physicians are in an excellent position to detect drug abuse, however, by exploring this history when confronted by a possible drug abuse–related clinical manifestation. Primary care physicians also can make significant contributions. Individuals with substance abuse–related medical conditions were more likely to remain abstinent when randomized to an integrated medical care and substance abuse treatment program.[1]

The active ingredients of brief intervention have been summarized by the acronym *FRAMES*: (1) *feedback* of personal risk or impairment (e.g., sharing abnormal test results, discussing medical complications), (2) emphasis on personal *responsibility* to change, (3) clear *advice* to change, (4) a *menu* of different options for change, (5) provider *empathy*, and (6) enhancement of patient *self-efficacy* or optimism. Physicians can refer to substance abuse treatment specialists; self-help groups (e.g., Narcotics Anonymous) are often part of a successful abstinence maintenance program.

Pharmacologic treatment of opioid abuse includes agonist, antagonist, mixed agonist-antagonist, or symptomatic treatment. With detoxification, the goal is amelioration of the symptoms of heroin or other opioid withdrawal by agonist substitution (e.g., methadone) or symptomatic treatment (e.g., clonidine). A new treatment involves the use of buprenorphine, which has agonist and antagonist properties. It has a better safety profile and produces less physical dependence. Hospitalized opioid-dependent patients may be treated with methadone for withdrawal symptoms by any physician. Methadone for the postdischarge treatment of opioid addiction is currently available only via specially licensed treatment facilities.

Prevention of relapse to active heroin abuse has been attempted most commonly by substitution of a safer drug (e.g., methadone, buprenorphine, or the second-line agent, *l*-acetyl-α-methadol) with similar pharmacologic properties to relieve the craving and withdrawal and to block some of the euphoric effects of heroin (Table 30–4). These medications (i.e., high-dose methadone [60 to 100 mg], buprenorphine [16 to 32 mg], and levomethadyl acetate [75 to 115 mg]) substantially reduced use of illicit opioids compared with low-dose methadone (20 mg) in a randomized controlled trial.[2]

Table 30–4 • RELAPSE PREVENTION FOR OPIOID ABUSE

MEDICATION	DOSE	DOSING INTERVAL	MECHANISM	PRESCRIBING REGULATIONS	WITHDRAWAL*
Methadone	60–100 mg orally	Daily	Agonist	Yes	++
LAAM	30–115 mg orally	q2–3d	Agonist	Yes	++
Buprenorphine	8–32 mg sublingually	q1–2d	Agonist/antagonist	Yes	+
Naltrexone	50 mg orally	Daily	Antagonist	No	–

LAAM = *l*-acetyl-α-methadol; FDA = Food and Drug Administration.
*++ = moderate; + = mild; – = none.

In an already detoxified patient, a less common alternative is to use an opioid antagonist (e.g., naltrexone) that effectively blocks agonist stimulation. Methadone is not adequate treatment for acute pain syndromes.

Although an emphasis on the treatment and prevention of drug abuse is crucial, physicians also can promote measures to reduce harm for injection drug users, including participation in needle exchange programs, avoidance of "shooting galleries" to obtain or administer drugs, prescriptions for needles and syringes, and instructions never to share "works" (injection equipment). These interventions, which can be delivered by physicians to drug abusers, have played a crucial role in international efforts to limit the spread of HIV infection.

Cocaine abuse is treated by psychotherapy, behavioral therapy, and 12-step programs. Acupuncture has been used for detoxification and for preventing relapse. As yet, no pharmacologic agent has been consistently effective in reducing cocaine use or craving. Dopamine agonists, antidepressants, and other drugs have been studied, but none are currently recommended. No antidote is known for acute cocaine overdose.

Marijuana use rarely requires acute treatment in the medical setting. Reassurance generally is sufficient to manage the occasional dysphoric manifestations. Occasionally, anxiety reactions require specific therapy with benzodiazepines; rarely, psychotic reactions are treated with haloperidol.

Specific therapy for the complications of *hallucinogen* use is nonpharmacologic and involves emotional reassurance and a calm supportive environment. No specific antagonists are clinically available for any of the hallucinogens. Medications are required only if the patient cannot be controlled adequately, in which case anxiolytic drugs are recommended.

Discontinuation of *benzodiazepines* can be accomplished in dependent patients by prescribing a regimen of gradual dose reduction. Alternatively, another long-acting sedative-hypnotic can be substituted for the drug of abuse and gradually withdrawn. It is important to attempt to verify that the patient has no alternative sources for these medications.

Future Directions

With current understanding of the associated morbidities and costs of drug use, increasing emphasis will be placed on the contributions that physicians can make in the care of patients with drug abuse and dependence. Opportunities to work with substance abuse providers to treat opioid-dependent patients in the primary medical care setting with pharmacologic therapy, including buprenorphine and methadone, will become widely available. Such future opportunities will increase the need for physicians to obtain skills to screen for drug abuse, address use of illicit drugs, and intervene to reduce the harm of these addictive behaviors.

Grade A
1. Weisner C, Mertens J, Parthasarathy S, et al: Integrating primary medical care with addiction treatment: A randomized controlled trial. JAMA 2001;286: 1715–1723.
2. 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.

SUGGESTED READINGS

Fiellin DA, O'Connor PG: Office-based treatment of opioid-dependent patients. N Engl J Med 2002;347:817–824. *A useful guide to ambulatory care.*

Kosten TR, O'Connor PG: Management of drug and alcohol withdrawal. N Engl J Med 2003;348:1786–1795. *Overview of inpatient and outpatient options.*

McLellan AT, Lewis DC, O'Brien CP, Kleber HD: Drug dependence, a chronic medical illness: Implications for treatment, insurance, and outcomes evaluation. JAMA 2000;284:1689–1695. *Drug dependence is portrayed as a mainstream medical problem, and its treatment is associated with efficacy and adherence comparable with other common chronic conditions.*

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. *A new paradigm of collaboration between physicians and substance abuse professionals is described, and potential benefits of such interactions are outlined.*

31 GLUCOCORTICOSTEROIDS IN RELATION TO INFLAMMATORY DISEASE

Paul Katz

For more than 50 years, glucocorticosteroids have been important agents in treating diseases characterized by inflammation and exaggerated immune responses. The pioneering work of Hench and colleagues in rheumatoid arthritis showed the possible potency of these agents in such pathologic states. Although substantial advances have been made in understanding the mechanisms by which glucocorticosteroids exert beneficial effects, considerable gaps in knowledge remain. Despite extensive data regarding the in vitro and in vivo activities of these drugs, it is probable that glucocorticosteroids have different beneficial activities in different diseases.

The challenge of glucocorticosteroid therapy continues to be the counterbalancing of desirable anti-inflammatory and immunosuppressive actions versus undesirable pharmacologic activities. More precise understanding of the mechanisms of action of glucocorticosteroids has not resulted in the development of regimens with minimal toxicity.

Pharmacology

The glucocorticosteroid preparations available for systemic use (Table 31–1) differ in their relative anti-inflammatory potency, potential for sodium retention, and plasma and biologic half-lives. In general, shorter acting preparations, such as prednisone and prednisolone, are preferable to longer acting agents, such as dexamethasone, because tapering to an alternate-day schedule cannot be accomplished with drugs with prolonged (i.e., >24 hours) biologic half-lives. Additionally, hydrocortisone and cortisone rarely are used to treat inflammatory and immunologically mediated diseases because of the considerable mineralocorticoid activity that accompanies their use.

Mechanisms of Action

Glucocorticosteroids exert anti-inflammatory and immunosuppressive actions through several pathways. Nonetheless, all effects

Table 31–1 • GLUCOCORTICOSTEROID PREPARATIONS

	ANTI-INFLAMMATORY POTENCY	EQUIVALENT DOSE (mg)	SODIUM-RETAINING POTENCY	PLASMA HALF-LIFE (min)	BIOLOGIC HALF-LIFE (h)
Hydrocortisone	1	20	2+	90	8–12
Prednisone	4	5	1+	60	12–36
Methylprednisolone	5	4	0	180	12–36
Betamethasone	20–30	0.6	0	100–300	36–54

From Garber EK, Targoff C, Paulus HE: In Paulus HE, Furst DE, Droomgoole SH (eds): Drugs for Rheumatic Diseases. New York, Churchill Livingstone, 1987, p 446.

0145-6008/04/2804-0572$03.00/0
ALCOHOLISM: CLINICAL AND EXPERIMENTAL RESEARCH

Vol. 28, No. 4
April 2004

Alcohol Consumption and Antiretroviral Adherence Among HIV-Infected Persons With Alcohol Problems

Jeffrey H. Samet, Nicholas J. Horton, Seville Meli, Kenneth A. Freedberg, and Anita Palepu

Background: Alcohol abuse has been associated with poor adherence to highly active antiretroviral therapy (HAART). We examined the relative importance of varying levels of alcohol consumption on adherence in HIV-infected patients with a history of alcohol problems.

Methods: We surveyed 349 HIV-infected persons with a history of alcohol problems at 6-month intervals. Of these subjects, 267 were taking HAART at one or more time periods during the 30-month follow-up period. Interviews assessed recent adherence to HAART and past month alcohol consumption, defined as "none", "moderate", and "at risk". We investigated the relationship between adherence to HAART and alcohol consumption at baseline and at each subsequent 6-month follow-up interval using multivariable longitudinal regression models, while controlling for potential confounders.

Results: Among the 267 HIV-infected persons with a history of alcohol problems who were receiving HAART, alcohol consumption was the most significant predictor of adherence ($p < 0.0001$), with better adherence being associated with recent abstinence from alcohol, compared with at-risk level usage (odds ratio = 3.6, 95% confidence interval = 2.1–6.2) or compared with moderate usage (odds ratio = 3.0, 95% confidence interval = 2.0–4.5).

Conclusions: Any alcohol use among HIV-infected persons with a history of alcohol problems is associated with worse HAART adherence. Addressing alcohol use in HIV-infected persons may improve antiretroviral adherence and ultimately clinical outcomes.

Key Words: HIV, Adherence, Alcohol, Highly Active Antiretroviral Therapy.

HIGHLY ACTIVE ANTIRETROVIRAL therapy (HAART) has led to substantial reductions in morbidity and mortality as well as improved quality of life for many HIV-infected individuals (Hogg et al., 1999; Palella et al., 1998; Revicki et al., 1999). Paterson et al. (2000) found that adherence to protease inhibitor therapy of 95% or greater is required for optimal HIV-RNA suppression. For most patients, however, actual adherence rates are often lower, with 40% to 60% of patients reporting <90% adherence (Bartlett, 2002). Reasons suggested by patients for not being fully adherent with HAART include forgetfulness, being away from home or too busy, or having a change in their daily routine (Bartlett, 2002). Other barriers to optimal adherence include psychiatric disorders such as depression or substance use (Arnsten et al., 2002; Lucas et al., 2001, 2002; Paterson et al., 2000; Starace et al., 2002), regimen complexity (Stone et al., 2001), and medication side effects (Ammassari et al., 2001).

Alcohol consumption is common among HIV-infected persons and has an important relationship to HAART adherence. The prevalence of heavy drinking in the National HIV Cost and Services Utilization Study among the 2864 HIV-infected patients in the United States in care was 8%. Factors significantly associated with heavy drinking were cocaine and heroin use, less education (<high school vs college), and not having a history of an AIDS-defining illness (Galvan et al., 2002). Cook et al. (2001) found that 19% of HIV-infected primary care patients reported problem drinking. Additionally, 33% consumed mild to moderate amounts of alcohol. Compared with nonproblem drinkers, problem drinkers were more likely to take their antiretroviral medications off schedule (Cook et al., 2001). Paterson et al. (2000) found a nonsignificant trend of fewer patients with alcohol problems among those with 95% or greater adherence compared with less than 95% adherence

From Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Department of Medicine (JHS, SM), Boston University School of Medicine; Department of Social and Behavioral Sciences (JHS), Boston University School of Public Health; Department of Mathematics (NJH), Smith College, Northampton, MA; Division of General Medicine and Partners AIDS Research Center (KAF), Massachusetts General Hospital, Boston, MA; and the Centre for Health Evaluation and Outcome Sciences (AP), St. Paul's Hospital, Department of Medicine, University of British Columbia, Vancouver, British Columbia.

Received for publication August 11, 2003; accepted December 16, 2003.

Supported, in part, by Grants RO1-AA13766 (Clinical Impact of HCV and Alcohol in HIV-Infected Persons), RO1-AA11785 (Medication Adherence in Alcohol Abusing HIV Patients), and RO1-AA10870 (Enhanced Linkage of Alcohol Abusers to Primary Care) from NIAAA; by USPHS Grant M01 RR00533; and by a Canadian Institute for Health Research New Investigator Award (AP).

Reprint requests: Jeffrey H. Samet, MD, MA, MPH, Boston Medical Center, 91 East Concord St., Suite 200, Boston, MA 02118; Fax: 617-414-4676; E-mail: jsamet@bu.edu.

DOI: 10.1097/01.ALC.0000122103.74491.78

(4% vs. 17%, $p = 0.13$). The relationship between alcohol and antiretroviral therapy adherence is important to investigate because alcohol consumption is a potentially modifiable characteristic associated with nonadherence (Cook et al., 2001).

In this study we further investigate the relationship between alcohol consumption and adherence to HAART. Our objective was to determine whether there is a safe level of alcohol consumption with regard to HAART adherence for HIV-infected patients with a history of alcohol problems. We examined these issues among participants in the HIV-Alcohol Longitudinal Cohort (HIV-ALC) study.

METHODS

Study Design

We analyzed data from a prospective cohort of HIV-infected patients with a history of alcohol problems. Subjects were observed from one to seven times over 36 months. One hundred and fifty-one subjects in the cohort participated in a randomized controlled trial of a HAART adherence intervention (ADHERE study; Samet et al., 2002); appropriate adjustments were made to the analysis to account for the trial.

Study Population

Patients who were HIV infected and had a history of alcohol problems were identified by explicit eligibility criteria. All potential subjects who gave two or more positive responses to the CAGE questionnaire (Ewing, 1984), a screening test for lifetime alcohol problems, were eligible. In addition, those patients recruited from the Boston Medical Center HIV Diagnostic Evaluation Unit (Samet et al., 1995) who did not meet CAGE criteria were eligible if one of two attending physicians (JS and KF) made a specific diagnosis of alcohol abuse or dependence. Other entry criteria included the following: fluency in English or Spanish, Mini-Mental State Examination score greater than 20 (Folstein et al., 1975), and no plans to move from the Boston area in the next 2 years. The Institutional Review Boards of both Boston Medical Center and Beth Israel Deaconess Medical Center approved the study.

From July 1997 through July 2001, recruitment of subjects occurred by multiple methods and from several sites: Boston Medical Center HIV Diagnostic Evaluation Unit 56%; posted flyers in the community 17%; Boston Medical Center Primary Care Clinic 13%; respite facility for homeless persons 5%; methadone clinic 4%; subject referrals 4%; and Beth Israel Deaconess Medical Center 2%. The eligibility criterion of a history of alcohol problems was determined by the CAGE questionnaire in 313 of 349 (90%) subjects and based on clinical assessment in 36 of 349 (10%) subjects. Diagnostic interviews for alcohol problems in a sample of these subjects ($n = 141$) revealed that 95% had a lifetime history of either alcohol dependence (80% [113 of 141]) or abuse (15% [21 of 141]; (Samet et al., 2004). Follow-up continued through July 2001.

Data Collection

After obtaining informed consent, a research associate interviewed subjects using a standardized instrument to ascertain baseline information including demographics, drug and alcohol consumption, use of substance abuse treatment services, use of HAART, and adherence to HAART in the preceding 6 months. We attempted to obtain CD4 cell counts and HIV RNA level on all subjects. Laboratory tests performed within 6 months of the interview as part of clinical care were recorded. If not available through routine clinical care, blood samples were obtained and tested for CD4 cell count and HIV RNA using the Boston Medical Center Clinical Laboratory. For the Spanish interview instrument, standardized scales in Spanish were used when available; the remainder of the questionnaire was translated from English into Spanish, back-translated to check for accuracy, and then corrected.

Outcome Variable

The outcome variable of interest was self-reported adherence to HAART. Using the AIDS Clinical Trials Group instrument, patients reported the names of their antiretroviral medications as well as the number of doses and the total number of pills prescribed daily (Chesney et al., 2000). The 3-day self-reported number of pills missed was computed for each HIV medication. We defined adherence as a dichotomous variable, in which patients less than 100% adherent over the prior 3 days were considered nonadherent.

Primary Independent Variable

Alcohol consumption was assessed by patient interview using a series of standardized questions on alcohol use. The battery included alcohol quantity and frequency questions and the Addiction Severity Index, an assessment instrument with well-documented reliability and validity, even among homeless persons (McLellan et al., 1985). To encourage truth telling about alcohol consumption, breath alcohol level was also measured before the interview (Gibb et al., 1984). We used the NIAAA guideline for at-risk drinking: greater than 14 drinks per week (or more than 4 drinks per day) for men, and greater than 7 drinks per week (or more than 3 drinks per day) for women (NIAAA Alcoholism, 1995). Alcohol consumption below these levels was considered moderate use in this study.

Other Variables

Other specific variables assessed included gender, age, race/ethnicity, perceived social support from friends (Procidano and Heller, 1983), homelessness (which was defined as having spent at least one night either on the street or in a shelter in the 6 months before the interview), depressive symptoms using the Center for Epidemiologic Studies Depression scale (CES-D; Andresen et al., 1994), and any history of AIDS-defining opportunistic infections, any heroin and cocaine use over the prior 30 days, CD4 cell counts, Log HIV RNA levels, number of medication doses per day, and involvement in the ADHERE study. Perceived social support was collected at the initial observation and all scheduled follow-up interviews except the second.

Analysis

Bivariate comparisons of drinking and adherence at initial assessment were conducted using analysis of variance for continuous variables and χ^2 test for categorical variables. Variables considered as covariates and included in the multivariable longitudinal logistic regression model were either previously demonstrated to be associated with adherence in the literature or perceived as clinically significant.

Multivariable logistic longitudinal regression analysis using generalized estimating equations (Liang and Zeger, 1986) was used to relate level of alcohol consumption (none, moderate, or at risk) to adherence to HAART over time. Because serial measures on the same individuals were considered for the medication adherence analyses, generalized estimating equations were used to adjust for correlation between these measures over time using an exchangeable working correlation matrix (Liang and Zeger, 1986; Zeger and Liang, 1986). Depending on the number of completed visits, each subject contributed from one to seven interviews. A main effect term of time was included in the model for each time point (6 degrees of freedom). We adjusted for the following predictors in our model: gender, race/ethnicity, age, social support, homelessness, use of cocaine, use of heroin, number of antiretroviral medication doses per day, depressive symptom score, and involvement in the ADHERE randomized controlled trial (intervention, control, and not enrolled). All of the predictor variables except for gender and age were allowed to vary with time. Because social support from friends was not collected at the second scheduled observation, the value from the initial observation was used at the first and

second time points. For the social support 14-item scale, we used all scale scores that had eight or more completed items. All analyses were carried out using SAS/STAT version 8.2 (SAS/STAT, 2001).

RESULTS

There were 267 subjects in this study who were taking HAART, of whom 215 (81%) were male; mean age was 41 years; two thirds were ethnic minorities. Subjects were interviewed a median of three times (range, one to seven visits over a maximum of 36 months). The number of subjects completing one to seven interviews is as follows: (1) 56; (2) 32; (3) 41; (4) 43; (5) 34; (6) 42; and (7) 19. Because some subjects were observed but did not respond to all questions, a total of 798 observations were included in the multivariable regression models. In a different analysis of data from the entire HIV-ALC cohort, Ehrenstein et al. (2004) found that time of recruitment into the study was the most important predictor of the number of completed interviews ($p < 0.0001$).

Other characteristics of the cohort at the initial observation are presented in Table 1. The mean CES-D score for depressive symptoms, 22, was above the threshold commonly used to indicate depressive symptoms (>16). One fourth ($n = 67$) reported being homeless in the prior 6 months. Heroin and cocaine use in the previous 30 days was reported by 10% and 25%, respectively. Within the 6 months before the initial observation, 60% of the cohort received substance abuse treatment services, 61% attended at least one self-help meeting per week, and 27% spent at least one night in jail. Sixty-six percent were fully adherent using 3-day self-report at the initial observation; mean number of doses of antiretroviral therapy per day was five.

The characteristics of the 205 subjects who were taking HAART at the initial observation, stratified by alcohol consumption, are shown in Table 2. The distribution of alcohol consumption was none (60%), moderate drinking (24%), and at-risk drinking (16%). Unadjusted comparisons at the initial observation indicate no significant differences in race/ethnicity, age, gender, social support, depressive symptoms, homelessness, and doses of antiretroviral therapy per day by level of alcohol consumption in the prior 30 days. Drinkers were more likely to use cocaine and heroin in the prior 30 days compared with nondrinkers. Additionally, worse adherence was associated with increased alcohol consumption, with abstainers being significantly more likely to report 100% adherence than moderate or at risk drinkers (76% vs. 57% vs. 42%, $p = 0.0004$).

Table 1. Characteristics of HIV-Infected Subjects With a History of Alcohol Problems Ever on HAART ($n = 267$)

Characteristic	*n* (%)
Female	52 (19)
Race/Ethnicity	
White	93 (35)
African American	112 (42)
Other	62 (23)
Heroin use, past month	26 (10)
Cocaine use, past month	67 (25)
Jail, past 6 months	73 (27)
Homeless, past 6 months	67 (25)
Any opportunistic infection, lifetime	79 (30)
Any substance abuse treatment, past 6 months	160 (60)
Self-help attendance, past 6 months	161 (61)
Taking HAART at initial observation	205 (77)
100% HAART 3-day adherence*	135 (66)
	Mean (SD)
Age, years	41.2 (7.4)
Social support friends**	9.3 (4.2)
Depressive symptoms (CES-D)	22.0 (12.8)
Doses of HAART/day***	5.0 (1.6)
Log HIV RNA level†	2.5 (2.0)
CD4 cell count‡	379 (255)

* $n = 204$, one subject missing 3-day adherence; ** Perceived Social Support Friend subscale, $n = 266$; *** for the 205 subjects on HAART at initial observation.
† $n = 259$; ‡ $n = 260$.

Among the 267 subjects, a total of 798 interviews were conducted during the study period in which subjects were receiving HAART and had complete data on all predictors. The results of the multivariable longitudinal regression model of factors associated with adherence are presented in Table 3. We found no association between the odds of being fully adherent and gender ($p = 0.80$), age ($p = 0.23$), race/ethnicity ($p = 0.55$), recent use of heroin ($p = 0.23$), recent use of cocaine ($p = 0.19$), and involvement in the ADHERE study ($p = 0.85$). There were no significant changes in adherence over time ($p = 0.97$). Subjects with higher CES-D scores (adjusted odds ratio [AOR] = 0.98, 95% confidence interval [CI] = 0.97–0.996) and more doses of HAART per day (AOR = 0.87, 95% CI = 0.78–0.97) had significantly worse adherence. Subjects with more social support from friends reported significantly better odds of being fully adherent (AOR = 1.06, 95% CI = 1.02–1.11). Alcohol consumption was an important predictor of adherence ($p < 0.0001$), with better odds of adherence being associated with recent abstinence from alcohol, compared with at-risk level usage (AOR = 3.6, 95% CI = 2.1–6.2). Abstainers also reported significantly better adherence compared with moderate users (AOR = 3.0, 95% CI = 2.0–4.5). There were no significant differences between moderate and at-risk use ($p = 0.47$).

DISCUSSION

Among our sample of HIV-infected persons with a history of alcohol problems, we found that alcohol was the most significant predictor of adherence, with better adherence being associated with recent abstinence from alcohol, compared with moderate or at-risk level usage. Although higher CES-D scores (more depressive symptoms) and increased doses of HAART per day were significantly associated with decreased adherence, alcohol consumption remained the most significant factor associated with adherence. Additionally, worse adherence to HAART occurred regardless of whether the alcohol consumption was at the moderate or at-risk level. Thus, among HIV-infected persons with a history of alcohol problems, there was no

Table 2. Characteristics of HIV-Infected Subjects With a History of Alcohol Problems Receiving HAART at Initial Observation (n = 205) Stratified by Level of Alcohol Consumption

	Abstinent 122 (60%)	Moderate 49 (24%)	At risk 33 (16%)	p Value
N (%)				
100% 3-day adherence*	93 (76)	28 (57)	14 (42)	0.0004
Female	26 (21)	8 (16)	7 (21)	0.76
Race/ethnicity				0.89
Black	47 (38)	22 (45)	15 (45)	
White	48 (39)	16 (33)	11 (33)	
Other	28 (23)	11 (22)	7 (21)	
Cocaine use, past month	5 (4)	19 (39)	24 (73)	<0.0001
Heroin use, past month	5 (4)	6 (12)	8 (24)	0.001
Homeless, past 6 months	28 (23)	12 (24)	6 (18)	0.79
Mean (SD)				
Age	40.9 (0.7)	41.5 (1.1)	40.5 (1.3)	0.82
Social support friends*	9.1 (0.4)	10.5 (0.6)	8.9 (0.7)	0.11
Depressive Symptoms (CES-D)	21.1 (1.1)	19.1 (1.7)	20.4 (2.1)	0.64
No. of doses per day*	5.1 (0.1)	4.9 (0.2)	5.1 (0.3)	0.90

* n = 204.

Table 3. Factors Associated With Adherence to HAART Among a Cohort of HIV-Infected Patients With a History of Alcohol Problems Based on Multivariable Logistic Regression (n = 798 Interviews in 267 Subjects)

Factor	3-day 100% adherence AOR (95% CI)
Alcohol consumption*	
None	3.6 (2.1–6.2)
Moderate	1.2 (0.7–2.0)
At-risk	1.0
Depressive symptom score=	0.98 (0.97–0.996)
Heroin**	0.7 (0.4–1.3)
Cocaine**	0.8 (0.5–1.4)
Social support friends$^{\varphi}$	1.06 (1.02–1.11)
Daily dose$^{\gamma}$	0.87 (0.78–0.97)
Age	1.02 (0.99–1.05)
Female vs. male	0.9 (0.5–1.6)
Race/ethnicity	
Black	0.8 (0.5–1.4)
White	1.0 (0.6–1.8)
Other	1.0

* Reference group is at-risk level usage of alcohol; = per 1 unit of the CES-D score; ** use in the past 30 days; φ per 1 unit of the Perceived Social Support Friend subscale, which ranges from 1 to 14; γ average no. of daily doses of antiretroviral medication.

safe level of drinking, as any amount of alcohol consumption was found to negatively affect adherence to HAART.

Alcohol consumption among persons with HIV has received increased attention due to the relevant behavioral and clinical issues that arise with respect to disease progression and adherence to HAART (Bagby et al., 2003; Miguez et al., 2003; Samet et al., 2003). Studies have suggested that the prevalence of alcohol abuse or dependence ranges from 20% to 40% among HIV-infected primary care patients (Cook et al., 2001; Lefevre et al., 1995). Cook et al. (2001) found that problem drinking was mildly associated with missed doses and significantly associated with taking medications off schedule. This is consistent with our findings that at-risk drinking and moderate drinking are associated with worse adherence. We also found significantly worse adherence among both moderate and at-risk drinkers in comparison to abstainers in contrast to Cook et al. (2001), who reported no significant difference in medication adherence among abstainers and mild-to-moderate drinkers. Another study of 140 HIV-infected patients at a county hospital found that alcohol use was independently associated with worse adherence. After adjusting for demographic and clinical factors, patients actively using drugs took 59% of doses vs 72% for nonusers, and those drinking alcohol took 66% of doses vs. 74% for nondrinkers (Golin et al., 2002).

One small (n = 94) study compared the virological outcomes of HIV-infected patients receiving HAART by their daily alcohol intake; they found no significant difference in the proportion of patients who achieved HIV-1 RNA suppression (Fabris et al., 2000). Another study of 220 HIV-infected drug users found that the prevalence of heavy alcohol consumption, defined as daily or three to four times per week, was 63%. Heavy alcohol users were more likely to be male and between 35 and 45 years of age. Compared with light or nondrinkers, heavy alcohol users receiving antiretroviral therapy were more likely to have CD4 counts below 500 cells/μl and less likely to achieve virological suppression (Miguez et al., 2003). Furthermore, given the prevalence of hepatitis C and HIV coinfection among drug users, those who use alcohol not only may be less likely to adhere to antiretroviral therapy for behavioral reasons but also may be less able to tolerate antiretroviral therapy due to hepatotoxicity (Sulkowski et al., 2000).

Depressive symptoms are prevalent among HIV-infected persons (Kilbourne et al., 2002) and have been reported to adversely affect adherence to HAART (Paterson et al., 2000; Turner et al., 2003), and this is consistent with our findings. Turner et al. (2003) studied factors associated with pharmacy-measured adherence to HAART in more than 5,000 drug users through the New York State Medicaid program. They found that women were less adherent than men and were more likely to be diagnosed with depression (34% vs. 29%). Our finding that an increased number of doses of therapy per day was negatively associated with adherence has also been noted elsewhere (Ammassari et al., 2002; Golin et al., 2002; Stone et al., 2001)

and highlights the need for simplified antiretroviral regimens.

Evidence for effective brief interventions among problem drinkers who are not HIV infected demonstrates that alcohol consumption should be perceived as a potential modifiable risk factor for poor adherence to HAART. Strategies for addressing alcohol use include presenting the negative consequences of alcohol use to the individual with risky drinking (Fleming et al., 1997; Ockene et al., 1999; Samet et al., 1996). Negative consequences of at-risk drinking and possibly any alcohol use among patients with a history of alcohol problems include poor HAART adherence.

This study has several limitations. Self-reported adherence tends to be overestimated, although in one study self-reported adherence among drug users was highly correlated with adherence as measured by medication event monitoring systems and was also correlated with HIV-1 RNA suppression (Arnsten et al., 2001). There is currently no gold standard for adherence, and although medication event monitoring systems may be more accurate than self-report, they can underestimate adherence and may be difficult to use in a clinical population as opposed to a medication trial (Samet et al. 2001).

In summary, we found that better adherence to HAART was associated with recent abstinence from alcohol compared with moderate or at-risk level usage among HIV-infected persons with alcohol problems. Alcohol consumption was the most significant factor associated with medication adherence in patients with alcohol problems. This study result is a significant advance in that it examines adherence to HAART in a detailed fashion in a population at high risk for poor adherence, namely those with alcohol problems. These findings also argue that with regard to adherence to HAART there may be no safe level of drinking among patients with a history of alcohol problems. In patients like these, alcohol consumption and medication adherence should both be closely monitored. The risk for suboptimal adherence to HAART when any amount of alcohol is used among patients with a history of alcohol problems should be explicitly discussed when reviewing the use of antiretroviral medications.

ACKNOWLEDGMENTS

Data management was provided by DM-STAT, Inc., Medford, MA, in particular Nicole Tibbetts, BA. We appreciate the contributions of the clinical staff of the HIV Diagnostic Evaluation Unit at Boston Medical Center, including Colleen LaBelle, RN and Jennifer Doyle. We also acknowledge contributions to this project by Danielle Appugliese and Carly Bridden. Finally, we appreciate the contributions of Howard Libman, MD, at the Beth Israel Deaconess Medical Center.

REFERENCES

Ammassari A, Murri R, Pezzotti P, Trotta MP, Ravasio L, De Longis P, Lo Caputo S, Narciso P, Pauluzzi S, Carosi G, Nappa S, Piano P, Izzo CM, Lichtner M, Rezza G, Monforte AD, Ippolito G, Moroni M, Wu AW, Antinori A (2001) Self-reported symptoms and medication side effects influence adherence to highly active antiretroviral therapy in persons with HIV infection. J AIDS 28:445–449.

Ammassari A, Trotta MP, Murri R, Castelli F, Narciso P, Noto P, Vecchiet J, D'Arminio Monforte A, Wu AW, Antinori A (2002) Correlates and predictors of adherence to highly active antiretroviral therapy: overview of published literature. J AIDS 31:S123–S127.

Andresen EM, Malmgren JA, Carter WB, Patrick DL (1994) Screening for depression in well older adults—evaluation of a short-form of the CES-D. Am J Prev Med 10:77–84.

Arnsten JH, Demas PA, Farzadegan H, Grant RW, Gourevitch MN, Chang CJ, Buono D, Eckholdt H, Howard AA, Schoenbaum EE (2001) Antiretroviral therapy adherence and viral suppression in HIV-infected drug users: comparison of self-report and electronic monitoring. Clin Infect Dis 33:1417–1423.

Arnsten JH, Demas PA, Farzadegan H, Grant RW, Gourevitch MN, Farzadegan H, Howard AA, Schoenbaum EE (2002) Impact of active drug use on antiretroviral therapy adherence and viral suppression in HIV-infected drug users. J Gen Intern Med 17:377–381.

Bagby GJ, Stoltz DA, Zhang P, Kolls JK, Brown J, Bohm RP, Rockar R, Purcell J, Murphey-Corb M, Nelson S (2003) The effect of chronic binge ethanol consumption on the primary stage of SIV infection in rhesus macaques. Alcohol Clin Exp Res 27:495–502.

Bartlett JA (2002) Addressing the challenges of adherence. J AIDS (Suppl 1) 29:S2–10.

Chesney MA, Ickovics JR, Chambers DB, Gifford AL, Neidig J, Zwickl B, Wu AW (2000) 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 12:255–266.

Cook RL, Sereika SM, Hunt SC, Woodward WC, Erlen JA, Conigliaro J (2001) Problem drinking and medication adherence among persons with HIV infection. J Gen Intern Med 16:83–88.

Ehrenstein V, Horton NJ, Samet JH (2004) Inconsistent condom use among HIV-infected patients with alcohol problems. Drug Alcohol Depend 73:159–166.

Ewing JA (1984) Detecting alcoholism. The CAGE questionnaire. JAMA 252:1905–1907.

Fabris P, Tositti G, Manfrin V, Giordani MT, Vaglia A, Cattelan AM, Carlotto A (2000) Does alcohol intake affect highly active antiretroviral therapy (HAART) response in HIV-positive patients? J AIDS 25:92–93.

Fleming MF, Barry KL, Manwell LB, Johnson K, London R (1997) Brief physician advice for problem alcohol drinkers—a randomized controlled trial in community-based primary care practices. JAMA 277: 1039–1045.

Folstein MF, Folstein SE, McHugh PR (1975) "Mini-mental state." A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 12:189–198.

Galvan FH, Bing EG, Fleishman JA, London AS, Caetano R, Burnam MA, Longshore D, Morton SC, Orlando M, Shapiro M (2002) 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 63:179–186.

Gibb KA, Johnston CC, Maring SD, Nowak RM (1984) Accuracy and usefulness of a breath analyzer. Ann Emerg Med 13:516–520.

Golin CE, Liu HH, Hays RD, Miller LG, Beck CK, Ickovics J, Kaplan AH, Wenger NS (2002) A prospective study of predictors of adherence to combination antiretroviral medication. J Gen Intern Med 17:756–765.

Hogg RS, Yip B, Kully C, Craib KJ, O'Shaughnessy MV, Schechter MT, Montaner JS (1999) Improved survival among HIV-infected patients after initiation of triple-drug antiretroviral regimens. Can Med Assoc J 160:659–665.

Kilbourne AM, Justice AC, Rollman BL, McGinnis KA, Rabeneck L, Weissman S, Smola S, Schultz R, Whittle J, Rodriguez-Barradas M

(2002) Clinical importance of HIV and depressive symptoms among veterans with HIV infection. J Gen Intern Med 17:512–520.

Lefevre F, O'Leary B, Moran M, Mossar M, Yarnold PR, Martin GJ, Glassroth J (1995) Alcohol consumption among HIV-infected patients. J Gen Intern Med 10:458–460.

Liang KY, Zeger SY (1986) Longitudinal data analysis using generalized linear models. Biometrika 73:13–22.

Lucas GM, Cheever LW, Chaisson RE, Moore RD (2001) Detrimental effects of continued illicit drug use on the treatment of HIV-1 infection. J AIDS 27:251–259.

Lucas GM, Gebo KA, Chaisson RE, Moore RD (2002) Longitudinal assessment of the effects of drug and alcohol abuse on HIV-1 treatment outcomes in an urban clinic. AIDS 16:767–774.

McLellan AT, Cassiola J, Griffith J, Evans F, Bass HL, O'Brien CP (1985) New data from the Addiction Severity Index: reliability and validity in three centers. J Nerv Ment Dis 173:412–423.

Miguez MJ, Shor-Posner G, Morales G, Rodriguez A, Burbano X (2003) HIV treatment in drug abusers: impact of alcohol use. Addict Biol 8:33–37.

NIAAA Alcoholism (1995) *The Physician's Guide to Helping Patients with Alcohol Problems.* National Institutes of Health, Washington, DC.

Ockene JK, Adams A, Hurley TG, Wheeler EV, Hebert JR (1999) Brief physician- and nurse practitioner-delivered counseling for high-risk drinkers—does it work? Arch Intern Med 159:2198–2205.

Palella FJ Jr, Delaney KM, Moorman AC, Loveless MO, Fuhrer J, Satten GA, Aschman DJ, Holmberg SD (1998) Declining morbidity and mortality among patients with advanced human immunodeficiency virus infection. HIV Outpatient Study Investigators. N Engl J Med 338:853–860.

Paterson DL, Swindells S, Mohr J, Brester M, Vergis EN, Squier C, Wagener MM, Singh N (2000) Adherence to protease inhibitor therapy and outcomes in patients with HIV infection. Ann Intern Med 133:21–30.

Procidano ME, Heller K (1983) Measures of perceived social support from friends and from family: three validation studies. Am J Community Psychol 11:1–24.

Revicki DA, Moyle G, Stellbrink HJ, Barker C (1999) Quality of life outcomes of combination zalcitabine-zidovudine, saquinavir-zidovudine, and saquinavir-zalcitabine-zidovudine therapy for HIV-infected adults with CD4 cell counts between 50 and 350 per cubic millimeter. PISCES (SV14604) Study Group. AIDS 13:851–858.

Samet JH, Horton NJ, Dukes K, Meli S, Tripp T, Labelle C, Freedberg KA (2002) Randomized controlled trial of a multi-dimensional intervention to enhance adherence to antiretroviral therapy in HIV-infected patients with a history of alcohol problems (abstract). Poster presented at: XIV International AIDS Conference, July 8, 2002, Barcelona.

Samet JH, Horton NJ, Traphagen E, Lyon S, Freedberg K (2003) Alcohol consumption and HIV disease progression: are they related? Alcohol Clin Exp Res 27:862–867.

Samet JH, Libman H, LaBelle C, Steger K, Lewis R, Craven DE, Freedberg KA (1995) A model clinic for the initial evaluation and establishment of primary care for persons infected with human immunodeficiency virus. Arch Intern Med 155:1629–1633.

Samet JH, Phillips SJ, Horton NJ, Traphagen ET, Freedberg KA (2004) Detecting alcohol problems in HIV-infected patients: use of the CAGE questionnaire. AIDS Res Hum Retroviruses 20:151–156.

Samet JH, Rollnick S, Barnes H (1996) Beyond CAGE: a brief clinical approach after detection of substance abuse. Arch Intern Med 156: 2287–2293.

Samet JH, Sullivan LM, Traphagen ET, Ickovics JR (2001) Measuring adherence among HIV-infected persons: is MEMS consummate technology? AIDS Behav 1:21–29.

SAS/STAT (2001) ®*Software: Changes and Enhancements, Release 8.2.* SAS Institute Inc., Cary, NC.

Starace F, Ammassari A, Trotta MP, Murri R, De Longis P, Izzo C, Scalzini A, D'Arminio Monforte A, Wu AW, Antinori A (2002) Depression is a risk factor for suboptimal adherence to highly active antiretroviral therapy. J AIDS 31:S136–S139.

Stone VE, Hogan JW, Schuman P, Rompalo AM, Howard AA, Korkontzelou C, Smith DK (2001) Antiretroviral regimen complexity, self-reported adherence, and HIV patients' understanding of their regimens: survey of women in the HER study. J AIDS 28:124–131.

Sulkowski MS, Thomas DL, Chaisson RE, Moore RD (2000) Hepatotoxicity associated with antiretroviral therapy in adults infected with human immunodeficiency virus and the role of hepatitis C or B virus infection. JAMA 283:74–80.

Turner BJ, Laine C, Cosler L, Hauck WW (2003) Relationship of gender, depression, and health care delivery with antiretroviral adherence in HIV-infected drug users. J Gen Intern Med 18:248–257.

Zeger SL, Liang KY (1986) Longitudinal data analysis for discrete and continuous outcomes. Biometrics 42:121–130.

AIDS RESEARCH AND HUMAN RETROVIRUSES
Volume 20, Number 2, 2004, pp. 151–155

Detecting Alcohol Problems in HIV-Infected Patients: Use of the CAGE Questionnaire

JEFFREY H. SAMET,[1,2] SHARON J. PHILLIPS,[1] NICHOLAS J. HORTON,[3] ELIZABETH T. TRAPHAGEN,[1] and KENNETH A. FREEDBERG[4]

ABSTRACT

The HIV epidemic has been consistently associated with injection drug use and crack cocaine, but alcohol problems in HIV-infected persons are less well described. Our objectives were 2-fold: (1) to assess the prevalence of alcohol problems in HIV-infected patients initiating medical care; and (2) to determine the positive predictive value of the CAGE questionnaire for alcohol abuse or dependence in HIV-infected patients. Between July 1997 and October 2000, we assessed a consecutive series of patients who were establishing primary care for HIV infection (clinic sample), using an established alcohol screening test, the CAGE questionnaire. In addition, we enrolled other HIV-infected patients, including some of the clinic sample, who had two or more positive responses to the CAGE questionnaire into a longitudinal cohort (cohort sample), performed a diagnostic interview for lifetime history of alcohol abuse and dependence, and determined the positive predictive value of CAGE for alcohol diagnoses. In the clinic sample (n = 664), 42% (276 of 664) had two or more positive responses to the four CAGE questions. In the cohort sample (n = 141), 95% (134 of 141) met *DSM-IV* criteria for diagnosis of lifetime alcohol abuse or dependence. For patients initiating HIV primary care, a history of alcohol problems is very common. The CAGE questionnaire identifies a lifetime history of alcohol abuse or dependence in HIV-infected patients. Routine screening for alcohol problems should be performed in all patients entering HIV medical care and the CAGE questions are useful in this setting.

INTRODUCTION

For the past two decades, focus on the association between HIV infection and substance abuse has been directed primarily at the injection of opioids and cocaine[1] and the use of crack cocaine.[2] Although alcohol is a more common substance of abuse in the United States than illicit drugs, its role in HIV infection and disease has received more limited attention. Because alcohol use influences HIV infection in several key ways, the role of screening for alcohol use disorders is an important clinical issue.

Alcohol use is known to be related to high-risk sexual behavior,[3] and alcohol dependence is a known risk factor for HIV infection, as seen in studies of patients in alcohol treatment units in New York and San Francisco, where 5–10% of clients had HIV infection, a prevalence not fully attributable to injection drug use.[4,5] Additionally, alcohol use has been associated with decreased adherence to antiretroviral medications[6,7] and increased clinical comorbidities of HIV infection.[8] Finally, alcohol use has been shown to be associated with HIV progression among patients on antiretroviral therapy.[9]

The prevalence of alcohol problems in general non-HIV primary care populations has consistently revealed that 20–30% of patients in primary care meet criteria for a diagnosis of past or present alcohol abuse or dependence.[10–12] Despite the high prevalence of alcohol disorders and the substantial burden of disease known to accompany them,[13,14] these problems are recognized by medical providers in less than half of affected patients.[12,15–17] The importance of screening for alcohol problems in primary care has been recognized by the U.S. Preventive Services Task Force and is included in its list of recommended activities.[18] The National Institute on Alcohol Abuse and Alcoholism (NIAAA) has set forth a guide for such interventions in routine clinical care.[19]

Although the need to address alcohol problems in primary care populations has been well documented, the importance of

Departments of [1]Medicine and [2]Social and Behavioral Sciences, Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Boston University Schools of Medicine and Public Health, Boston, Massachusetts 02118.
[3]Department of Mathematics, Smith College, Northampton, Massachusetts 01063.
[4]Division of General Medicine and Partners AIDS Research Center, Massachusetts General Hospital, Boston, Massachusetts 02114.

assessing alcohol use among HIV-infected patients has been underappreciated. As the significance of alcohol use in HIV-infected patients becomes increasingly apparent, most notably in the context of antiretroviral medication adherence[20,21] and risky sexual behaviors,[3,22] this lack of attention to alcohol issues becomes more problematic.

The CAGE questionnaire, an alcohol screening instrument recommended by NIAAA, has been examined extensively for use in primary care populations.[10,23] It is used to assess a history of alcohol abuse or dependence. In primary care populations, two or more positive responses to the CAGE questionnaire have documented sensitivity in the range of 73–82% and specificity of approximately 90%.[10] Its ease of administration and test characteristics make it a potentially valuable instrument for use among HIV-infected persons. In this study we characterize the prevalence of alcohol problems in HIV-infected patients initiating primary medical care and examine the use of the CAGE questionnaire as a screening tool in this population.

MATERIALS AND METHODS

Study design

We studied the prevalence of alcohol problems in a consecutive series of HIV-infected patients at the time of initial physician assessment (clinic sample). In a separate cohort (cohort sample), which included some of the clinic sample, we conducted an extensive assessment of subjects, including clinical diagnosis of alcohol abuse or dependence, as part of a longitudinal study. This study was approved by the Institutional Review Board of the Boston Medical Center.

Study populations and assessment

The clinic sample. We assessed all HIV-infected patients initiating HIV-related primary care at Boston Medical Center between July 1997 and October 2000 in the HIV Diagnostic Evaluation Unit (DEU). The DEU was an intake and assessment clinic for HIV-infected persons without a physician in order to link them to primary care. Ongoing primary care was not provided at this weekly clinic site.[24] An attending physician used the CAGE questionnaire at the first visit to screen patients fluent in English or Spanish for alcohol problems as part of routine clinical care. The CAGE questionnaire is comprised of the following four questions:

Have you ever felt you should **C**ut down on your drinking?
Have people **A**nnoyed you by criticizing your drinking?
Have you ever felt bad or **G**uilty about your drinking?
Have you ever had a drink first thing in the morning to steady your nerves or get rid of a hangover (**E**ye opener)?[25,26]

Some studies recommend a cutoff of one or more positive responses to raise suspicion of the presence of an alcohol problem,[19,25] while others advocate a cutoff of two or more for a higher degree of specificity.[26] We applied the more strict cutoff of two or more positive responses.

The cohort sample. We assessed study participants enrolled in the longitudinal cohort study recruited from either the DEU or other Boston Medical Center Primary Care Clinics between July 1997 and July 2001. In addition to HIV infection and two positive responses to CAGE questions, entry criteria for the cohort study included no definite plans to move from the study area in the next 2 years and a Mini-Mental State Examination[27] score of at least 21 (scale 0 to 30). After informed consent was obtained, study subjects underwent research interviews.

Subjects enrolled in the longitudinal cohort study underwent multiple interviews. Demographic characteristics, primary HIV risk factor, self-reported liver problems, clinical laboratory values, current alcohol consumption, and alcohol-related diagnoses by the alcohol section of the Composite International Diagnostic Interview (CIDI) were obtained.[28] The CIDI is a standardized, structured interview developed collaboratively by an international team of researchers with the purpose of creating an epidemiological instrument to provide standardized diagnoses of psychiatric disorders, including alcohol dependence and abuse as defined by *DSM-IV* criteria.[29] The instrument has been shown to be reliable in many different cultures and reduces discrepancies among interviewers.[30] This test has been used to generate a lifetime diagnosis of alcohol dependence or abuse. Trained interviewers administered the CIDI alcohol module in the context of this research study.

Alcohol consumption was assessed by patient interview using a series of standardized questions on alcohol use. The battery included alcohol quantity and frequency questions and the Addiction Severity Index, an assessment instrument with well-documented reliability and validity, even among homeless persons.[31] To encourage truth telling about alcohol consumption, breath alcohol level was also measured.[32] From this interview we determined if the subject was abstinent ("none"), or met criteria for "moderate" or "at-risk" drinking, using the National Institute on Alcohol Abuse and Alcoholism guideline for at-risk drinking of greater than 14 drinks per week (or more than four drinks per day) for men, and greater than seven drinks per week (or more than three drinks per day) for women.[33] Alcohol consumption below these levels was considered moderate use in this study.

Self-report of health information was solicited; subjects were asked if a doctor had ever told them that they had hepatitis or liver disease, and if they were currently taking highly active antiretroviral therapy (HAART). We attempted to obtain CD4 cell counts and HIV RNA levels on all subjects. Laboratory tests performed within 6 months of the interview as part of clinical care were recorded. If not available through routine clinical care, blood samples were obtained and tested for CD4 cell count and HIV RNA using the Boston Medical Center Laboratory.

Research associates interviewed subjects in English or Spanish. The Spanish interview instrument used the standardized Spanish versions of scales when available. The remainder of the Spanish questionnaire was translated from the English version, back-translated to check for accuracy, and then corrected.

Analysis

Analyses were performed using PC SAS statistical software, version 8. Positive predictive value and exact 95% confidence intervals were estimated by determining the proportion of CAGE-positive patients who tested positive on the CIDI.

RESULTS

The clinic sample

Of the 755 patients evaluated at the clinic during the study period, 715 (95%) spoke English or Spanish and are described in Table 1. Seventy percent of the sample were men and 77% were ethnic minorities. The most common HIV risk factor was injection drug use (47%); heterosexual sex was reported by 37% and men having sex with men (MSM) was reported by 16%. The average age of subjects was 39 years. Of these eligible patients, 93% (664 of 715) were evaluated for alcohol problems with the CAGE questionnaire: 42% (276 of 664) (95% CI 38%, 45%) had two or more positive CAGE responses.

Based on data from all patients administered the CAGE in the clinic, alcohol problems were more common in injection drug users (55% prevalence), however, they were also very common in patients with other risk behaviors. Among MSM, 30% had a positive CAGE, as did 30% of those with heterosexual risk.

The cohort sample

The characteristics of the 141 patients in the cohort sample are outlined in Table 2. Eighty-five percent of patients were men and 74% were ethnic minorities. The most common HIV risk factor was injection drug use (65%); heterosexual sex was reported by 23% and MSM by 12%. The average age of subjects was 41 years. Subjects did not differ by demographics, substance use, or HIV disease markers based on site of enrollment.

Of the 141 patients in the cohort sample, 95% (95% CI 90%, 98%) met *DSM-IV* criteria for diagnosis of either lifetime alcohol abuse or dependence; 113 of 141 (80%) met criteria for dependence and 21 of 141 (15%) met criteria for abuse; 7 of 141 (5%) did not meet criteria for either dependence or abuse. Prevalence of a *DSM-IV* diagnosis of alcohol abuse or dependence among CAGE-positive patients was 98% (88 of 90) in patients with risk behavior injection drug use, and 90% (46 of 51) in patients with other risk behaviors. While diagnoses for current alcohol abuse or dependence were not obtained, 60 (43%) subjects reported current alcohol consumption: 16% of the cohort sample reported current at-risk drinking and an additional 27% current moderate drinking.

Table 1. Characteristics of HIV-Infected Patients Seen at the HIV Diagnostic Evaluation Unit (DEU) between July 1997 and October 2000 and Eligible for Assessment for Alcohol Problems (n = 715)

Characteristic	n *(%)*
Gender	
Male	498 (70%)
Female	217 (30%)
Race	
Black	357 (50%)
Latino	179 (25%)
White	163 (23%)
Other	16 (2%)
Primary HIV risk factor[a]	
Injection drug use	334 (47%)
Heterosexual	266 (37%)
Men who have sex with men	112 (16%)
Age (mean, years)	39

[a]Information available on n = 712.

Table 2. Characteristics of HIV-Infected Patients Enrolled in Longitudinal Cohort Study Between July 1997 and July 2001 (n = 141)

Characteristic	n *(%)*
Gender	
Male	120 (85%)
Female	21 (15%)
Race	
Black	68 (48%)
White	37 (26%)
Other	36 (26%)
Primary HIV risk factor	
Injection drug use	92 (65%)
Heterosexual	32 (23%)
Men who have sex with men	17 (12%)
Age (mean, years)	41
CIDI *DSM-IV* lifetime alcohol diagnosis	
No diagnosis	7 (5%)
Abuse	21 (15%)
Dependence	113 (80%)
Current drinking[a]	
Abstinent	80 (57%)
Moderate	38 (27%)
At risk	22 (16%)
Current use of HAART	77 (55%)
Liver disease[b]	92 (66%)
Hepatitis[b]	49 (35%)

[a]Information available on n = 140.
[b]Self-report.

DISCUSSION

Although injection drug use dominates substance abuse discussions regarding the AIDS epidemic, appreciation of the role of alcohol in HIV infection is increasing. Alcohol use is an important patient characteristic related to several key HIV-related clinical issues including prevention strategies,[34] linkage to medical care,[35] and the use of antiretroviral therapy.[6] Antiretroviral usage may be limited due to interactions between alcohol and certain medications, including didanosine, zalcitabine, and stavudine. In addition, alcohol consumption is associated with decreased adherence to antiretroviral therapy and other treatment regimens,[36–38] the effectiveness of which is predicated upon an extremely high level of adherence.[39]

In addition to limiting treatment options, alcohol use may disrupt other aspects of HIV care. Diseases aggravated by alcohol use often overlap those associated with HIV, including hepatitis, pancreatitis, peripheral neuropathy, and bacterial pneumonia.[8] Excessive alcohol use may be an important exacerbating factor in the clinical course of an individual with such comorbidities. Hepatitis C, a common coinfection with HIV, is

adversely affected by excessive alcohol use, resulting in more rapid and frequent progression to cirrhosis.[40]

The CAGE questionnaire, a standard test for the presence of a lifetime history of alcohol abuse or dependence, detected a substantial number (42%) of HIV-infected patients with two or more positive CAGE responses in the clinic sample. Alcohol problems are more common in this population than the 20–30% prevalence estimated in non-HIV primary care settings.[11,12] These findings are consistent with those found in a smaller sample of 111 mostly white HIV-infected patients from a university hospital in the Midwestern United States in which 41% were categorized as having alcoholism according to the Michigan Alcohol Screening Test.[41]

The CAGE questionnaire, a simple, inexpensive screening test that takes less than 1 minute to administer, yielded a positive predictive value of 95% for a lifetime diagnosis of alcohol problems. This result is consistent with test characteristics previously reported for the CAGE questionnaire in other settings, with studies publishing a sensitivity of 74–75%, specificity of 91–96%, and positive predictive value of 82% for the CAGE, using a two-positive response cutoff.[10,42] The higher positive predictive value in this study is likely due to a greater prevalence of alcohol problems in this population than in the outpatient populations surveyed in earlier studies (estimated at 36% and 20%, respectively).

The CAGE screening test detects a lifetime history of alcohol problems. We also found that 16% of those interviewed reported at-risk drinking in the prior 30 days, and 27% reported moderate drinking in the prior 30 days. Addressing alcohol use in these patients is crucial because of its impact on many aspects of HIV care. In addition, knowledge of a prior history of an alcohol problem is an important component of the clinical history, since recognition of the success of people in recovery from alcohol problems is described as an essential aspect of clinical care.[43,44]

This study has several limitations. We are not able to examine the sensitivity or specificity of the alcohol screening instrument, since patients who did not meet the CAGE two or more criteria were not administered the full diagnostic interview. Also, as the CAGE screening test is reportedly 70–80% sensitive, the 42% prevalence of alcohol problems in the clinic sample is likely an underestimation. Another limitation is the study's reliance on the clinic sample drawn from a single urban clinic that is not necessarily representative of all HIV-infected persons in the United States. A disproportionately high number of subjects had injection drug use as an HIV risk behavior, and alcohol use is common in drug users. However, the presence of a lifetime history of alcohol abuse or dependence was not found solely in those HIV-infected patients with a history of injection drug use. Alcohol problems were found commonly in all HIV-infected patient subgroups in this urban hospital.

Many concerns demand attention in the HIV patient initiating medical care, including clinical, social, and family issues. However, the prevalence of alcohol problems found in this study coupled with the alcohol-associated clinical problems described in the medical literature strongly suggests that alcohol screening merits inclusion in the standard initial assessment of all HIV-infected patients.[45] With at least 95% adherence with antiretroviral therapy as the goal for all treated HIV-infected patients,[39] physician knowledge of a patient's history of alcohol use is imperative. The CAGE test accurately identifies patients with alcohol problems and requires a minimal time commitment on the part of clinicians. Alcohol is clearly the "other substance" affecting the lives of HIV-infected patients, and routine screening for alcohol problems for all patients initiating medical care for HIV infection should be done.

ACKNOWLEDGMENTS

The authors appreciate the contributions of the clinical staff of the HIV Diagnostic Evaluation Unit at Boston Medical Center (Colleen LaBelle, RN, Jennifer Doyle, Sandra Looby-Gordon, MD, and Margaret Sullivan, MD), staff researchers on the project (Vera Doljanskaia, Kathleen Doyle, and Katie Schoolwerth), manuscript editors (Seville Meli and Emily Williams), and data management assistance (Tara Tripp and Nicole Tibbetts) of DM-STAT. Support for this study came from the following grants from the National Institute on Alcohol Abuse and Alcoholism (NIAAA) of the NIH: RO1-AA13766 (Clinical Impact of HCV and Alcohol in HIV-Infected Persons), RO1-AA11785 (Medication Adherence in Alcohol Abusing HIV Patients), and RO1-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, USPHS Grant M01 RR00533.

REFERENCES

1. O'Connor PG, Selwyn PA, and Schottenfeld RS: Medical care for injection-drug users with human immunodeficiency virus infection. N Engl J Med 1994;331:450–459.
2. Booth RE, Watters JK, and Chitwood DD: HIV risk-related sex behaviors among injection drug users, crack smokers, and injection drug users who smoke crack. Am J Public Health 1993;83: 1144–1148.
3. Stall R, McKusick L, Wiley J, Coates TJ, and Ostrow DG: Alcohol and drug use during sexual activity and compliance with safe sex guidelines for AIDS: The AIDS Behavioral Research Project. Health Educ Quart 1996;13:359–371.
4. Mahler J, Yi D, Sacks M, *et al.*: Undetected HIV infection among patients admitted to an alcohol rehabilitation unit. Am J Psychiatry 1994;151:439–440.
5. Avins AL, Woods WJ, Lindan CP, Hudes ES, Clark W, and Hulley SB: HIV infection and risk behaviors among heterosexuals in alcohol treatment programs. JAMA 1994;271:515–518.
6. Whitfield RM, Bechtel LM, and Starich GH: The impact of ethanol and marinol/marijuana usage on HIV+/AIDS patients undergoing azidothymidine, azidothymidine/dideoxycytidine, or dideoxyinosine therapy. Alcohol Clin Exp Res 1997;21:122–127.
7. Galvan FH, Bing EG, Fleishman JA, *et al.*: The prevalence of alcohol consumption and heavy drinking among people with HIV in the United States: Results from the HIV Cost and Service Utilization Study. J Stud Alcohol 202;63:179–186.
8. Samet JH: Alcohol and HIV infection. Alcoscope 2000;3:2–4.
9. Samet JH, Horton NJ, Traphagen ET, Lyon SM, and Freedberg KA: Alcohol consumption and HIV disease progression: Are they related? Alcohol Clin Exp Res 2003;27:862–867.
10. Buchsbaum DG, Buchanan G, Centor RM, Schnoll SH, and Lawton MJ: Screening for alcohol abuse using CAGE scores and likelihood ratios. Ann Intern Med 1991;115:774–777.

11. Fleming MF, Manwell LB, Barry KL, and Johnson K: At-risk drinking in an HMO primary care sample: Prevalence and health policy implications. Am J Public Health 1998;88:90–93.
12. Cleary PD, Miller M, Bush BT, Warburg MM, Delbanco TL, and Aronson MD: Prevalence and recognition of alcohol abuse in a primary care population. Am J Med 1988;85:466–471.
13. National Institute on Alcohol Abuse and Alcoholism: 10th Special Report to the U.S. Congress on Alcohol and Health. NIH Pub. No. 00-1583, 2000.
14. Michaud CM, Murray CJ, and Bloom BR: Burden of disease—implications for future research. JAMA 2001;285:535–539.
15. Rydon R, Redman S, Sanson-Fisher RW, and Reid AL: Detection of alcohol-related problems in general practice. J Stud Alcohol 1992;53:197–202.
16. Moore RD, Bone LR, Geller G, Mamon JA, Stokes EJ, and Levine DM: Prevalence, detection, and treatment of alcoholism in hospitalized patients. JAMA 1989;261:403–407.
17. Coulehan JL, Zettler-Segal M, Block M, McClelland M, and Schulberg HC: Recognition of alcoholism and substance abuse in primary care patients. Arch Intern Med 1987;147:349–352.
18. U.S. Preventive Services Task Force: *Guide to Clinical Preventive Services: A Report of the U.S. Preventive Services Task Force*. Williams & Wilkins, Baltimore, 1996.
19. National Institute on Alcohol Abuse and Alcoholism: *The Physicians' Guide to Helping Patients With Alcohol Problems*. National Institute on Alcohol Abuse and Alcoholism, Bethesda, MD, 1995. NIH Pub. No. 95-3769.
20. Eldred LJ, Wu AW, Chaisson RE, and Moore RD: Adherence to antiretroviral and pneumocystis prophylaxis in HIV disease. J Acquir Immune Defic Syndr Hum Retrovirol 1998;18:117–125.
21. Cook RL, Sereika SM, Hunt SC, *et al.*: Problem drinking and medication adherence among persons with HIV infection. J Gen Intern Med 2001;16:83–88.
22. Fitterling JM, Matens PB, Scotti JR, and Allen JS: AIDS risk behaviors and knowledge among heterosexual alcoholics and non-injecting drug users. Addiction 1993;88:1257–1265.
23. Saitz R, Lepore MF, Sullivan LM, Amaro H, and Samet JH: Alcohol abuse and dependence in Latinos living in the United States: Validation of the CAGE (4M) questions. Arch Intern Med 1999;159:718–724.
24. Samet JH, Libman H, LaBelle C, *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;155:1629–1633.
25. Ewing JA: Detecting alcoholism: The CAGE questionnaire. JAMA 1984;252:1905–1907.
26. Mayfield D, McLeod G, and Hall P: The CAGE questionnaire: Validation of a new alcoholism screening instrument. Am J Psychiatry 1974;131:1121–1123.
27. Folstein MF, Folstein SE, and McHugh PR: Mini-mental state. A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12:189–198.
28. World Health Organization: *Composite International Diagnostic Interview (CIDI)*. World Health Organization, Geneva, Switzerland, 1996.
29. American Psychiatric Association: *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*. American Psychiatric Association, Washington, D.C. 1994.
30. Robins LN, Wing J, Wittchen HU, *et al.*: The composite international diagnostic interview: An epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. Arch Gen Psychiatry 1988;45:1069–1077.
31. McLellan AT, Kushner H, Metzger D, *et al.*: The fifth edition of the Addiction Severity Index. J Subst Abuse Treat 1992;9:199–213.
32. Gibb KA, Yee AS, Johnston CC, Martin SD, and Nowak RM: Accuracy and usefulness of a breath alcohol analyzer. Ann Emerg Med 1984;13:516–520.
33. National Institute on Alcohol Abuse and Alcoholism: *The Physician's Guide to Helping Patients with Alcohol Problems*. National Institutes of Health, Washington, D.C., 1995. NIH Publ. No. 95-3769.
34. Boscarino JA, Avins AL, Woods WJ, Lindan CP, Hudes ES, and Clark W: Alcohol-related risk factors associated with HIV infection among patients entering alcoholism treatment: Implications for prevention. J Stud Alcohol 1995;56:642–653.
35. Samet JH, Freedberg KA, Stein MD, *et al.*: Trillion virion delay: Time from testing positive for HIV to presentation for primary care. Arch Intern Med 1998;158:734–740.
36. Vanhove GF, Schapiro JM, Winters MA, Merigan TC, and Blaschke TF: Patient compliance and drug failure in protease inhibitor monotherapy. JAMA 1996;276:1955–1956.
37. Shafer RW, Winters MA, Palmer S, and Merigan TC: Multiple concurrent reverse transcriptase and protease mutations and multidrug resistance of HIV-1 isolates from heavily treated patients. Ann Intern Med 1998;128:906–911.
38. Kastrissios H, Suárez J-R, Katzenstein D, Girard P, Sheiner LB, and Blaschke TF: Characterizing patterns of drug-taking behavior with a multiple drug regimen in an AIDS clinical trial. AIDS 1998;12:2295–2303.
39. Paterson DL, Swindells S, Mohr J, *et al.*: Adherence to protease inhibitor therapy and outcomes in patients with HIV infection. Ann Intern Med 2000;133:21–30.
40. Pol S, Lamorthe B, Thi NT, *et al.*: Retrospective analysis of the impact of HIV infection and alcohol use on chronic hepatitis C in a large cohort of drug users. J Hepatol 1998;28:945–950.
41. Lefevre F, O'Leary B, Moran M, *et al.*: Alcohol consumption among HIV-infected patients. J Gen Intern Med 1995;10:458–460.
42. Bush B, Shaw S, Cleary P, Delbanco TL, and Aronson MD: Screening for alcohol abuse using the CAGE questionnaire. Am J Med 1987;82:231–235.
43. Samet JH, Rollnick S, and Barnes H: Beyond CAGE: A brief clinical approach after detection of substance abuse. Arch Intern Med 1996;156:2287–2293.
44. Friedmann R, Saitz R, and Samet JH: Relapse prevention in primary care: Management of adults recovering from alcohol or other drug problems. JAMA 1998;279:1227–1231.
45. Gallant JE: The seropositive patient—the initial encounter: *Medscape. HIV/AIDS Clinical Management Modules*, 2001.

Address reprint requests to:
Jeffrey H. Samet
Clinical Addiction Research and Education (CARE) Unit
Section of General Internal Medicine
91 East Concord St., Suite 200
Boston Medical Center
Boston, Massachusetts 02118

E-mail: jsamet@bu.edu

Establishment of a Multidisciplinary Health Evaluation and Linkage to Primary Care (HELP) Clinic in a Detoxification Unit

Luann P. Sweeney, RN
Jeffrey H. Samet, MD, MA, MPH
Mary Jo Larson, PhD
Richard Saitz, MD, MPH

ABSTRACT. We evaluated the feasibility of establishing a multidisciplinary Health Evaluation and Linkage to Primary care (HELP) clinic at an urban residential detoxification unit. Patients received a clinical eval-

Luann P. Sweeney is affiliated with the Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Department of Medicine, Boston Medical Center and Boston University School of Medicine, Boston, MA.

Jeffrey H. Samet is affiliated with the Departments of Social and Behavioral Sciences and Richard Saitz is affiliated with the Department of Epidemiology, Boston University School of Public Health, Boston, MA. Jeffrey H. Samet and Richard Saitz are also affiliated with the CARE Unit, Boston Medical Center.

Mary Jo Larson is affiliated with the New England Research Institutes, Watertown, MA.

Address correspondence to: Richard Saitz, MD, MPH, Boston Medical Center, 91 East Concord Street #200, Boston, MA 02118 (E-mail: rsaitz@bu.edu).

This work was supported by grants from the National Institute on Alcohol Abuse and Alcoholism (R01-AA 10870) and the National Institute on Drug Abuse (R01-DA 10019). Dr. Saitz was Generalist Physician Faculty Scholar of the Robert Wood Johnson Foundation during the time this study was done. This work was supported in part by General Clinical Research Center grant MO1 RR00533 from the National Center for Research Resources.

Preliminary results were presented at the Association for Medical Education and Research on Substance Abuse annual national meeting, November 3, 2000.

Journal of Addictive Diseases, Vol. 23(2) 2004
http://www.haworthpress.com/web/JAD

Digital Object Identifier: 10.1300/J069v23n02_03

uation and facilitated linkage to primary medical care including personalized referral, reminders, and appointment rescheduling. Of 235 adults reporting alcohol, cocaine or heroin as first or second drug of choice and without a primary care physician, 178 (76%) received a full HELP clinic evaluation, 35 (15%) some clinic components, and 7 (3%) only a primary care appointment. Of those with a full evaluation, 28% received pneumococcal vaccination, and most received health behavior counseling. Over the subsequent 2 years, 131 (60%) of the 220 patients whom had any contact with the HELP clinic had at least one primary care visit. A multidisciplinary health clinic to evaluate patients during detoxification is feasible and can link patients with substance dependence to primary medical care.

KEYWORDS. Alcoholism, alcohol dependence, drug dependence, primary care, multidisciplinary team, integrated care, linkage

INTRODUCTION

Abuse of drugs and alcohol is associated with a large burden of morbidity and mortality,[1] the increased expense of emergency care,[2] in-patient hospitalizations and transmission of HIV infection.[3] Liver disease, accidents and injury, cancers, cardiovascular disease and psychiatric illness are among the co-morbidities associated with substance abuse.[4] Linkage of substance abusers to primary medical care may be able to reduce or at least address these consequences.[5] Primary medical care (hereafter referred to as primary care), is, according to the Institute of Medicine, "the provision of integrated, accessible health care services by clinicians who are accountable for addressing a large majority of personal health care needs, developing a sustained partnership with patients, and practicing in the context of family and community."[6] But despite an apparent increased need for medical care, those with substance abuse problems do not receive adequate medical care.[7-9] Saitz et al. assessed 5,824 persons presenting for substance abuse treatment in a public substance abuse treatment system and found that 41% did not have a physician.[9]

One model for providing primary care to this population is a distributive model in which patients in substance abuse treatment are referred to existing medical care.[4,10-14] This model is less expensive, likely more appealing to patients in recovery, and more easily replicable in the existing US health care system than the integrative approach (e.g., medical care provided on-site at a

specialty substance abuse treatment program).[5] However, the integrative approach has the advantage of bringing care to the patients. Primary care for substance abusers should be comprehensive and continuous; it should contribute to preventing or minimizing complications, providing early treatment of medical problems and assisting in the care of substance abuse issues.[4]

This paper describes the establishment of a novel mixed "integrative-distributive" approach to link patients undergoing residential detoxification for alcohol and drug problems to primary medical care. This approach, a multidisciplinary medical clinic in a freestanding, urban detoxification center, was called the Health Evaluation and Linkage to Primary care (HELP) clinic. The purpose of the clinic was to involve inpatients at the substance abuse detoxification facility in a single, initial, comprehensive health evaluation and then arrange subsequent follow-up with a primary care physician and ongoing health care.

METHODS

This is a descriptive report of the development of a novel, multidisciplinary clinic in a detoxification unit, and of cohort of patients evaluated therein. The HELP clinic was designed based on a comparable hospital-based clinic, the HIV Diagnostic Evaluation Unit[15] where HIV-infected patients without primary care are assessed by a multidisciplinary team, and linked with a primary care clinician. The three main clinical purposes of the HELP clinic were to provide for the patient: (1) multidisciplinary assessment; (2) education that underscored the importance of establishing primary medical care; and (3) referral for continuing medical care with a primary care clinician.

Patients

Patients evaluated in the HELP clinic were subjects randomized to the intervention group in a clinical trial. Enrollment details are reported elsewhere.[16] All subjects were recruited at least 24 hours after admission to a freestanding residential detoxification unit in Boston, Massachusetts between June 1, 1997 and April 1, 1999 and had the following characteristics: (1) alcohol, heroin or cocaine as a first or second drug of choice; (2) age over 17 years; (3) lived or planned to live in the Boston Medical Center (BMC) catchment area (the site at which primary care was to be arranged) or homelessness.

Exclusions were: (1) an established primary care relationship that the individual intended to continue; (2) mental status deficiencies; (3) specific plans to leave the Boston area in the next 12 months; (4) inability to provide three contact names; (5) pregnancy; and (6) not fluent in English or Spanish. Subjects (a maximum of 4 per session) were assigned to attend the next available

clinic session after enrollment. Trained staff researchers interviewed subjects to assess demographics, drug use, and history of preventive care testing (HIV test, Pap smear), as part of a more extensive research interview.[16]

HELP Clinic Data Collection, Outcomes and Analysis

A HELP clinic log recorded all phone calls, letters, and face-to-face contacts with patients after the clinic, and documented all time spent in 5-minute increments. The log also recorded primary care appointments scheduled, whether or not they were kept, and patient reasons for missed appointments. Patients signed a separate consent form for vaccination. This form was used to track receipt of vaccines. At the end of patient recruitment we reviewed HELP clinic records to identify clinic assessments and interventions. Completion of patient primary care appointments was tracked for two years using administrative data (the hospital's appointment scheduling system), patient contact by telephone and/or mail, and contact with primary care offices outside of BMC by the social worker. Descriptive statistics (mean ± SD for continuous variables, and proportions for dichotomous variables) are used to present results.

RESULTS

Implementation of the HELP Clinic

The HELP clinic met two afternoons each week in dedicated, sparsely equipped space within the confines of the detoxification unit. This detoxification unit was freestanding, not for profit, and space and clinical amenities were limited. Although a physician provided standing detoxification orders, no comprehensive or ongoing medical services were routinely provided. Counselors and nurses focused on detoxification services, there was no social worker to provide either concrete or therapeutic services. The clinical team included a physician trained in internal medicine [authors JHS, RS, and 3 others], a registered nurse [LPS], and a social worker. Each member of the HELP team was trained in a full day session in the principles of motivational interviewing by one of the study investigators with prior training as trainers in these methods [JHS, RS].[17] The session consisted primarily of role play of patients with substance dependence with the goal of primary care linkage, after a didactic description of the principles of motivational interviewing and advice to read the relevant text.[17] Each member of the HELP team used brief motivational counseling approaches to address substance dependence and linkage with primary medical care.

Each clinician was allocated 30 minutes with each patient. The HELP clinic nurse located patients randomized to the clinic, many of whom were sleeping

or reluctant to participate, and escorted them to the clinic. The nurse obtained a standardized health history including a review of the patient's medical, surgical and physical trauma history. Patients were asked about testicular or breast self-exam (TSE/BSE), tuberculosis infection, medications used, number of meals each day, amount of fat in the diet, consumption of fruits and vegetables, any regular exercise routine, amount of regular walking or physical labor, use of condoms and contraceptives, HIV testing, and knowledge and practice of needle sharing, bleaching or exchange. They were asked about family history of cancer, heart disease and other illness, involvement as a victim, witness or perpetrator of interpersonal violence as a child or adult; psychiatric hospitalization, outpatient treatment and history of or presence of homicidal/suicidal attempts or ideation. Those with current suicide or homicide risk were immediately referred to the HELP clinic social worker, who would refer to the detoxification unit counseling staff. Vital signs, height, and weight were measured. Patients were offered TSE/BSE "shower cards" (plastic cards which could be hung where one would be reminded frequently and which gave visual and verbal instructions in English and Spanish), HIV and needle cleansing written materials and condoms. Patients were offered and given pneumococcal vaccination if they had not previously had it. The nurse highlighted issues of concern for physician attention.

The physician then reviewed and expanded upon the medical history and performed a physical examination. Medical issues requiring immediate attention were referred to a nearby urgent care center. Identified medical issues were discussed with the patient, with particular emphasis on those of interest to the patient related to addiction and/or requiring ongoing medical care. Laboratory testing was not done routinely, both to minimize the cost of the clinic, and because these tests could be included as part of a recommended plan for the receiving primary care physician. Urine dipstick and stool guaiac testing were done when indicated by symptoms.

The social worker then obtained the patient's psychosocial history including demographics necessary to register and schedule a primary care appointment, review of government benefits, employment, housing, marital and parenting status, substance abuse, treatment and recovery history, discharge plans, psychiatric history, legal issues involving child support, criminal involvement, immigration issues and history of primary medical care. Much of this interview focused on concrete needs (food stamps, housing, legal issues) and available resources were offered to the patient, including referrals and phone calls made on the patient's behalf.

Aside from notifying detoxification unit staff of any urgent medical issues (e.g., excessive sedation, hypotension, urinary tract infection) the HELP clinic staff did not play a role in the detoxification per se, which included treatment

with long acting benzodiazepines, buprenorphine, methadone and/or other symptomatic treatments as appropriate.

Scheduling Primary Care Appointments. Primary care appointments were generally scheduled by the social worker at the adult medicine primary care clinic of a medical center, staffed by over 100 resident and attending physicians (including the HELP clinic physicians) and nurse practitioners. This inner city academic center clinic saw 15,000 unique patients in 45,000 visits annually during the study period (regardless of ability to pay).

In order to avoid the usual 3-month wait, the primary care clinic reserved a block of appointments for the HELP clinic patients. Appointments were scheduled with an attending physician or with a first or second year resident physician in a 3-year internal medicine training program. Patients provided preferences about the gender of the physician, time of appointment, language (English or Spanish) or specific physician if they knew one, which were generally accommodated. Dates of appointments were assigned with attention to patient discharge plans. Often patients were going on for further substance abuse treatment; in such scenarios they were often not allowed to attend any outside appointments from two weeks to three months. Patients could also choose specialty primary care clinics: a Latino Clinic with all Spanish speaking staff, a clinic with staff familiar with the specific needs of homeless people, and a clinic for those with HIV infection. Patients were sometimes able to schedule appointments with the same physician seen at the HELP clinic, and those with any interest in methadone treatment were offered an appointment with a primary care physician who was also a HELP clinic physician and a methadone clinic medical director. The patient was then given an appointment card with the name and number of the HELP clinic social worker as well as the name, phone number, appointment date, time and location of the new primary care physician. When a patient had a strong preference for care at another facility (a rare occurrence), the appointment was scheduled there. Patients going on for further addiction treatment were given a letter to take with them about the appointment.

Partial Exposure to the HELP Clinic. Since the clinic was twice a week, not daily, if a patient planned to leave the detoxification program before the HELP clinic, the nurse or social worker attempted to assess the individual's needs, discuss the value of establishing primary medical care, obtain medical and psychosocial history and arrange a primary care appointment. At times no contact was possible or only a primary care appointment was provided.

Multidisciplinary Referral. The team sent a physician-dictated referral letter to the primary care clinician based on the HELP clinic assessment including pertinent history, exam findings and specific recommendations for follow-up medical and addiction care.

Follow-Up. The social worker served in a case management role to facilitate linkage of patients to primary care. There was no limit to the number of primary care appointments made; active efforts continued for one year or until the appointment was kept. This role included patient contact by phone, mail and face-to-face contacts after discharge from the detoxification unit. The social worker also attempted to contact the patient's family and friends, if necessary. The face-to-face contact occurred in both unscheduled and scheduled meetings at the medical center or if the subject returned to the detoxification program. Patients were contacted through family, friends, community services, homeless shelters, treatment facilities, and other institutions. Letters included reminders prior to appointments, letters for missed appointments and a holiday letter to those who had not kept a primary care appointment.

Patient Participation in the HELP Clinic. Of the 235 individuals assigned to attend the HELP clinic, 220 (94%) received some component of the HELP clinic. Most (178/235, 76%) attended the clinic as designed, which included the health history, physical exam, psychosocial evaluation and primary care appointment scheduling. Some (35/235, 15%) received only social work or nursing evaluation and primary care appointment scheduling. A few (7/220, 3%) received only a primary care appointment and 15/235 (6%) did not receive any of the clinic components, having left the detoxification program between the time of assignment and the next scheduled clinic; three of this latter group scheduled and kept a primary care appointment themselves.

Patients and Outcomes

Of the 220 patients who received some component of the HELP clinic, 75% were male, 47% African-American, 38% stated cocaine, and 33% alcohol as their 1st drug of choice (Table 1). The majority (57%) was employed during the previous year, but more than half (55%) of patients earned less than $20,000 in any one of the previous 5 years. A third (33%) of the patients spent at least a night in jail during the previous 6 months and 47% had spent at least one night in a homeless shelter or on the street in the previous 6 months. Although all patients were without a primary care relationship, the majority had obtained some routine health screening. Of the women in the study, 96% had a Pap smear in the past 3 years and most patients (73%) had been tested for HIV in the past.

Almost all 178 patients attending a HELP clinic had health behaviors addressed by assessment and counseling by clinic staff (Table 2). A significant minority (28%) received pneumococcal vaccination. A total of 401 appointments were scheduled for the 220 patients who had any contact with the HELP clinic. Of those 220 scheduled for at least one primary care appointment, 100% were verified by clinic staff; 131/220 (60%) linked to primary care in

TABLE 1. Characteristics of patients who received Health Evaluation and Linkage to Primary care (HELP) clinic services (n = 220).

Characteristic	% (n)
Male	75 (166)
Race	
African-American	47 (104)
White	33 (73)
Hispanic	11 (26)
Other	8 (17)
Drug of choice	
Cocaine	38 (83)
Alcohol	33 (72)
Heroin	29 (64)
Homeless*	47 (104)
Unemployed	45 (103)
Income < $20,000 in past year	55 (125)
Incarcerated	
Past 6 months	33 (72)
Past 5 Years	53 (124)
Ever had a Pap smear	96 (52/54 women)
Ever been HIV tested	73 (161)
Ever used injection drugs	34 (74)

*> 1 night in a shelter or on the street in the previous 6 months
Pap = Papanicolaou smear for cervical cancer screening

the subsequent 2 years. The majority 110/131 (84%) of the patients who linked kept either the first or second appointment and 91/128 (71%) did so within 60 days of discharge from the detoxification unit (for 3, the date of the appointment was unknown). The mean days from detoxification to primary care appointment kept were 58.5 days (range 1 day to 405 days). From 1 to 7 appointments were scheduled per patient (mean 1.68). Most (70%) were with the general primary care clinic, 21% were with the homeless clinic and much smaller numbers were in the other specialized primary care clinics, including 9 (2%) who chose to have appointments made at another facility. Of the appointments that were not kept (n = 270), most (63%) were no shows; the patient canceled 67 (25%), and 32 (12%) were canceled by the practice. Of the 96 patients who reported reasons for missing 156 of these appointments, the most common reasons were: primary care appointment scheduling and office prob-

TABLE 2. Receipt of health assessment and counseling by HELP clinic attendees (n =178).*

Topic	% (n)
Testicular/breast self exam	93 (167)
Domestic violence	92 (158)
Condom use	99 (175)
HIV testing	100 (178)
Injection drug use	100 (178)
Dietary habits	94 (167)
Exercise habits	89 (158)
Tobacco use	100 (178)

*Applies to the 178 patients who had a full clinic evaluation (81% of the 220 who received any HELP clinic services).

lems, childcare and family obligations, conflicts with employment, conflicts with being in treatment, and substance use. Case management efforts to link patients with primary care were greater for those who never linked with primary care; patients who linked required a median of 7 phone calls, 1 letter, and 45 minutes of HELP clinic staff effort after the initial HELP clinic visit (Table 3).

DISCUSSION AND CONCLUSIONS

As hypothesized, we found that it is feasible to establish a clinic to begin to address the medical and psychosocial needs of patients with addictions while they are in residential detoxification by multidisciplinary assessment and referral to primary medical care. As reported elsewhere, almost half of the subjects in this study had chronic medical problems such as hypertension and asthma.[18]

We also found that patients referred attended such a clinic, and most completed the 90-minute session. The majority of these patients without primary medical care received health risk behavior counseling, a significant proportion received the indicated pneumococcal vaccination, and 60% linked with primary medical care. And efforts to facilitate linkage beyond the initial clinic were modest. Patients able to specify why they missed a scheduled primary care appointment most commonly reported the following reasons: employment, being in addiction treatment, using alcohol or drugs, legal issues, and primary care office scheduling.

TABLE 3. Case management efforts to link 220 subjects who received any HELP clinic component with primary care.

For subjects who linked with primary care

Effort	N*	Mean (SD)	Median (Range)
Phone calls	121	8.6 (7.3)	7 (1-36)
Letters	100	1.8 (1.3)	1 (1-7)
Time spent (minutes)**	125	74.0 (70.8)	45 (5-465)

For subjects who *never* linked with primary care

Effort	N*	Mean (SD)	Median (Range)
Phone calls	78	13.2 (9.4)	11 (1-49)
Letters	81	3.3 (2.2)	3 (1-11)
Time spent (minutes)**	85	101.4 (83.5)	85 (5-400)

*N is the number of patients receiving any of the specified effort.
**Time spent in phone calls, writing and mailing letters or talking face to face after the first HELP clinic visit.

This study had several strengths as a test of feasibility. First, the clinic was based on a previously successful model.[15] Second, the clinic was brought to patients rather than having patients come to the clinic, and it was done at a reachable moment–during a residential detoxification. Third, the clinic staff used standardized forms and approaches to patient evaluation and data recording, and they were trained in motivational interviewing, implementing this counseling style as they addressed linkage to primary care with patients. Lastly, for linkage, the follow-up was 100% since appointments records were derived from the administrative data at one medical center (and a small number elsewhere that were confirmed individually by telephone).

The study also has limitations. The primary limitation is generalizability. The applicability of our findings may be limited to uninsured urban populations without primary medical care at residential detoxification units. However, if effective in such a challenging population, it suggests efficacy if more broadly applied. Any replication of this clinic will require linkage with a primary care site that cares for patients regardless of ability to pay. Another limitation that should be considered in interpreting our results is the fact that subjects were compensated to enroll in the randomized trial in which a research intervention (HELP clinic) occurred. However, they were not specifically compensated to attend the clinic, and in fact, almost one quarter did not receive all clinic components, and some received none. Lastly, caution should

be used in drawing conclusions regarding the efficacy of the clinic because this report is of a prospective cohort, not the results of a randomized trial. But during the HELP clinic, health behavior counseling and vaccination that would not have occurred at all on-site at the detoxification unit was delivered. In addition, the high proportion of inner city adults with addictions, without primary medical care, and with markers indicating low socioeconomic status who in fact linked with primary care is remarkable, and higher than expected.

We are not aware of reports of examples of any other mixed "integrative-distributive" models for linking patients with addictions to primary medical care. The current US health care system primarily relies on a distributive model, in which detoxification patients are referred to medical care elsewhere. With this predominant model, patients with addictions often do not receive comprehensive, longitudinal primary care; rather they use medical care episodically for urgent problems.[19,20] The integrative model, or "one-stop shopping" in which all services (e.g., medical and addiction) are delivered on site, has been tried. In one randomized trial, integrated care was associated with increased abstinence in patients who had substance abuse-related medical conditions.[21] In another, veterans with alcohol dependence who received integrated care were more likely to be abstinent.[13] And although high rates of service delivery have been associated with this method,[22,23] the model has not been widely adopted, perhaps due to cost concerns (need to duplicate primary care delivery at addiction treatment units), and concerns about continuing primary care at a treatment unit after the patient is in a more stable recovery. Our approach combines the convenience of on-site initial delivery, with integration into the existing primary care system once linked. We speculate that the cost of implementing our approach would be between the integrative and distributive approaches since it requires on-site staff and space, but not for all medical care. This approach, if proven efficacious, may have a greater chance at widespread implementation since it requires less change in the health system, likely will cost less, and borrows features from integrative approaches that are responsible for efficacy, at least in the short-term. A randomized clinical trial of implementation of the HELP clinic addressed the question of efficacy, showing that linkage can be increased.[16] As a result, policymakers may wish to consider establishing HELP clinics in detoxification units. The fact that many patients with addictions do not have primary medical care at all suggests interventions are needed.

In addition to the demonstration of feasibility and its implications, our data provide some insight into reasons why patients with addictions miss primary care appointments. Prior studies addressing barriers to medical care for patients in addictions treatment have identified transportation as a key barrier[24,25] and program and health system characteristics.[8] In this single system, several reasons for missed appointments (in addition to relapse) surfaced as

important considerations when designing strategies to link patients with primary care. As with the general population, employment can be a barrier to attending a medical visit. We speculate that in this population, patients of lower socioeconomic status, and those with work consequences from their addictions may have even greater difficulty than the general population in taking time away from a job to see the doctor. Similarly, patients with addictions may have greater difficulty rescheduling and negotiating the system when a physician's office cancels or reschedules. A barrier that was unanticipated deserves particular attention: substance abuse treatment. In treatment, soon after detoxification, many programs restricted contact that patients could have with the "outside world." These restrictions applied to medical appointments (unless they were urgent). These barriers would all certainly be addressable in future interventions.

We conclude that many patients with addictions who do not have primary medical care, but who do have significant medical needs, can be identified and can have a medical and psychosocial evaluation and initial preventive interventions when they are admitted for detoxification. Such an intervention has been shown to improve linkage with primary medical care.[16] It remains to be determined whether addressing barriers we identified will improve linkage, and whether linkage itself will improve health and health care utilization.

REFERENCES

1. Michaud CM, Murray CJ, Bloom BR. Burden of disease–implications for future research. JAMA. 2001; 285:535-539.

2. McGeary KA, French MT. Illicit drug use and emergency room utilization. Health Serv Res. 2000; 25:153-169.

3. Centers of Disease Control and Prevention. HIV/AIDS Surveillance Report. 2000; 12:31-35.

4. Samet JH, Saitz R, Larson MJ. A case for enhanced linkage of substance abusers to primary medical care. Subst Abuse. 1996; 17:181-192.

5. 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.

6. Donaldson MS, Yordy KD, Lohr KN, Vanselow NA, eds. Primary Care: America's Health in a New Era. Washington, DC: National Academy Press, 1996.

7. D'Aunno T, Vaughn TE. An organizational analysis of service patterns in outpatient drug abuse treatment units. J Subst Abuse. 1995; 7:27-42.

8. D'Aunno T. Linking substance abuse treatment and primary health care. In Eerton, JA, Fox DM, Leshner AI, eds. Treating Drug Abusers Effectively. London: Basil Blackwell, 1997:311-351.

9. Saitz R, Mulvey KP, Samet JH. The substance-abusing patient and primary care: Linking via the addiction treatment system? Subst Abuse. 1997; 18:187-195.

10. Selwyn PA, Feingold AR, Iezza A, et al. Primary care for patients with human immunodeficiency virus (HIV) in a methadone maintenance program. Ann Intern Med. 1989; 110:761-763.

11. O'Connor PG, Molde S, Henry SP, et al. Human immunodeficiency virus infection in intravenous drug users: a model for primary care. Am J Med. 1992; 93:382-386.

12. Umbricht-Schneiter A, Ginn DH, Pabst KM, Begelow GE. Providing medical care to methadone clinic patients: referral vs. on-site care. Am J Pub Health. 1994; 84:207-210.

13. Willenbring ML, Olson DH. A randomized trial of integrated outpatient treatment for medically ill alcoholic men. Arch Intern Med. 1999; 159:1946-1952.

14. Schlenger WE, Kroutil LA, Roland EJ. National Evaluation Models for Linking Drug Abuse Treatment and Primary Care: A Casebook. Report prepared for the National Institute on Drug Abuse under contract number 283-90-001. Research Triangle Park, Research Triangle Institute, 1992.

15. Samet JH, Libman H, LaBelle C, 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; 155:1629-1633.

16. Samet JH, Larson MJ, Horton NJ, et al. Linking alcohol and drug dependent adults to primary medical care: a randomized controlled trial of a multidisciplinary health evaluation in a detoxification unit (the Health Evaluation and Linkage to Primary care [HELP] study). Addiction. 2003; 98:509-516.

17. Miller WR, Rollnick S. Motivational Interviewing. Preparing People to Change Addictive Behavior. New York: The Guilford Press, 1991.

18. DeAlba I, Samet JH, Saitz R. Burden of medical illness in drug and alcohol dependent persons without primary care. Am J Addict. 2004;13 (in press).

19. Bodenheimer T. Disease management in the American market. BMJ. 2000; 320:563-566.

20. Kimball HR, Young PR. Statement on the generalist physicians from the American Boards of Family Practice and Internal Medicine. JAMA. 1994; 271:315-316.

21. Weisner C, Mertens J, Parthasarathy S, Moore C, Lu Y. Integrating primary medical care with addiction treatment: a randomized controlled trial. JAMA. 2001; 286:1715-1723.

22. O'Connor PG, Farren CK, Rousanville BJ, O'Malley SS. A preliminary investigation of the management of alcohol dependence with naltrexone by primary care providers. Am J Med. 1997; 103:477-482.

23. Selwyn PA, Bunder NW, Wasserman WC, Arno PS. Utilization of on-site primary care services by HIV-seropositive and seronegative drug users in a methadone maintenance program. Pub Health Rep. 1993; 108:492-495.

24. Friedmann PD, D'Aunno TA, Jin L, Alexander JA. Medical and psychosocial services in drug abuse treatment: do stronger linkage promote client utilization? Health Serv Res. 2000; 35:443-465.

25. 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.

Managing the Community Response to Bioterrorist Threats

Crisis Health Risk Self-Assessment Tools to Triage the Patient Surge

VICTOR W. WEEDN, MICHAEL D. McDONALD, STEVEN E. LOCKE, MERRITT SCHREIBER, ROBERT H. FRIEDMAN, RICHARD G. NEWELL, AND LYDIA R. TEMOSHOK

Addressing the psychological and behavioral aftereffects of bioterrorism has emerged as a central but relatively neglected component of the National Bioterrorism Response effort. The significant public health challenges surrounding the potential psychosocial effects of bioterrorism were highlighted in the aftermath of 9/11; the anthrax attacks in Florida, New York, and Washington during 2001; national reports from the Institute of Medicine, Secretary's Emergency Public Information and Communication Advisory Board, and the National Advisory Board for Children and Terrorism; and compounded by results of several national exercises such as TOPOFF 2 and Dark Winter.

The psychosocial impacts of bioterrorist events are diverse and pervasive, extending far beyond those victims who are exposed and ill to include those individuals who perceive themselves or their family members as threatened or affected by the agent. Psychosocial effects are manifest across a spectrum of disorders, including posttraumatic stress disorder, panic, specific fears, and "psychogenic illness" with symptoms that may mimic those associated with various chemical, biological, radiological, or nuclear explosive (CBRNE) agents [1]. These phenomena have been described as multiple unexplained physical symptoms (MUPS) [2], idiopathic psychosomatic illness, disaster somatization reaction [3] and, in the past, as "worried well," a serious misnomer given that the overwhelming anxiety of such individuals hardly constitutes "wellness." When such reactions occur in large numbers of people, the phenomenon has been called "mass psychogenic illness" [4].

Experience with CBRNE incidents to date has shown that the great proportion of individuals seeking care (the so-called "surge") have in fact had no known direct exposure whatsoever but have been stimulated by media reports. For example, in the aftermath of the Tokyo sarin attack of 1995, approximately 88% of the medical emergency visits (4,500 of the 5,100) were of individuals who feared that they had been exposed but for whom no exposure was determined [5]. Similarly, following the death of four children who accidentally ingested discarded medical isotopes, over 100,000 individuals were screened, and approximately 5,000 with no known exposure to the agent in question reported symptoms consistent with acute radiological poisoning [6]. Thus, in terrorist events with CBRNE agents, the numbers of individuals who seek care in the absence of exposure may be orders of magnitude larger than the number of individuals seeking care who are actually exposed or ill, thus presenting a secondary public health impact and demand on emergency care systems beyond that of the CBRNE agent itself.

The healthcare response required by acutely ill patients in need of emergency medical care and/or hospitalization could be seriously compromised by this patient surge [7]. In other words, MUPS patients—and an even greater number of concerned but nonexposed individuals—threaten to overwhelm the system and in turn create further secondary mass psychological reactions, potentially delaying diagnosis and care for those actually exposed. From both the emergency management and public health standpoints, responses must be developed to address this phenomenon. Although the impact of surge has been postulated and observed in national exercises, the management of the psychosocial and behavioral factors that contribute to surge has been largely overlooked in most terrorism preparedness and response planning. To date, no national emergency management and public health strategy has been articulated to prevent, mitigate, and manage these psychological/behavioral reactions as they relate to medical surge following terrorism events.

Proposed Management of Patient Surge: Automated Crisis Health Risk Self-Assessment Tools

We will describe an armamentarium of automated crisis health risk self-assessment tools designed to address the CBRNE patient surge and form the basis for a community crisis awareness system. Harnessing the power of information technology (IT) yields the opportunity to reach the broad public in a sufficiently timely and efficient manner to deal with the overwhelming numbers of individuals who are likely to seek healthcare and information in the wake of bioterrorism. IT architectures can screen large numbers of individuals rapidly, efficiently, and uniformly; can apply sophisticated algorithms and heuristics to the problem of differential diagnosis; and can adapt to evolving information and situations on a real-time basis. We envision this IT screening system as only one component of a larger strategy to respond comprehensively to the various medical, psychological, and behavioral aspects of bioterrorism. When appropriately validated, the crisis health

 0739-5175/04/$20.00©2004IEEE

Terrorist biological weapons have proven to be more of a psychological threat than a physical danger.

risk self-assessment tools could be expanded into a cost-effective "all-hazards approach" to inform and educate the public about their risks during various phases of such emergencies.

Specifically, we propose developing, adapting, and validating automated crisis health risk self-assessment tools, consisting of a screening tool to determine the likelihood of actual exposure and whether the symptoms appear to be more likely psychogenic or agent-based, and a personalized report to help the individual determine the likelihood of actual exposure and need for further information or human services. These tools will provide individuals with the capacity to assess their own symptoms and relative risk of illness from a particular event prior to or upon entry into the healthcare system, thus diverting a portion of the nonexposed away from emergency and primary care settings towards other support systems and mental healthcare. One section of their personal report will provide referral information so that an individual can rapidly act on his or her most appropriate response options. The toolset can be designed to be customized and localized so that local jurisdictions can refer the public in a highly specific manner to the most optimal local services to address their particular medical, psychosocial, mental health, and/or informational service needs. The system could also assist in scheduling needed appointments. Thus, intelligent health risk self-assessment tools will facilitate the distribution of patients across and within healthcare systems geographically and temporally, permitting public health systems and communities to more appropriately and effectively handle the "surge," thus preventing regional healthcare provider systems from being overwhelmed.

Particularly in the case of radiological and biological events, perceptions of risk and individuals' reactions to a terrorist attack will change over time. Such projected trajectories will be factored into the proposed system's algorithms on a real-time, as-needed basis. We envision a mature system composed of multiple stages of assessment over time, designed to take into consideration special populations with special needs, such as children, the disabled, the disenfranchised, the elderly, or other populations with special healthcare needs.

Automated crisis health risk self-assessment tools could be deployed to homes, workplaces, schools, hospitals, or other community facilities through home computers, telephones, manned terminals, or some combination of these. The most obvious, easiest, and most powerful medium would be a Web site on the Internet accessible to individuals in their home. Most individuals have telephones (98% of Americans), and the Web-based version could be mirrored on the phone system. National and local trusted communicators would increase the value and receptivity of these communications. Individuals who do not have computers or telephones could be directed through the media and word of mouth to public-access stations. Manned or unmanned health risk self-assessment stations could be set up at emergency room entrances, schools, or other disaster healthcare settings for those not remaining at home.

Moreover, an intelligent crisis health risk self-assessment toolset could be incorporated into a highly interactive and personalized approach on the Web and through telephone-based interactive voice response (IVR) to provide additional value beyond present static media sources. For example, if the self-assessment results suggest that a person with dizziness has a likelihood of a psychogenic disorder, the program could provide the following feedback:

> Ms. Smith, given your answers to this self-assessment, it is most likely that you have not been exposed but are understandably nervous about the possibility of contracting <anthrax>. Sometimes nervousness can cause symptoms of dizziness, and we suggest you take the following anxiety relieving techniques ...

An example of such an integrated assessment and interactive educational feedback system was developed in the early 1990s by the Behavioral Prevention Program of the U.S. Military's HIV Research Program. Interactive videodisk systems for at-risk and HIV-infected military personnel consisted of: 1) self-paced assessment of, respectively, HIV/STD exposure and transmission risk, and 2) feedback modules, in which an individual's responses to a risk scenario triggered presentations of the consequences of various choices and/or corrective information [8]. Also, randomized clinical trials of automated telephone patient assistance systems have demonstrated clinical efficacy and reduction in emergency health services utilization by patients [9]–[14]. Evaluations by patients, consumers, and health professionals have indicated high levels of satisfaction with using IVR patient assistance programs such as the telephone-linked communications (TLC) discussed below.

Basis for Discriminant Functions

There is a long and distinguished history of the use of psychological testing dating back to World War II, with more recent developments in intelligent medical diagnostic and triage algorithms. A number of assessment tools, both self-report questionnaires and structured interviews, have been validated specifically for identifying and diagnosing a variety of more or less relevant psychological reactions and conditions (Table 1) [15], [16]. For example, "Impact of Event Scale-Revised" (IES-R), is a 22-item self-report measure designed to assess current subjective distress for any specific life event [17]–[19]. Recently, a National Institute for Mental Health (NIMH) roundtable compared the sensitivity and specificity of dozens of such instruments with unpublished results at this time [20].

These currently available instruments are generally narrow in their scope of assessment and have not been designed to distinguish psychiatric morbidity and psychosocial distress from acute physical illness. Stress and anxiety indicators, in and of themselves, are not expected to discriminate well between exposed and nonexposed populations. In the immediate aftermath of mass casualty events and the resultant media exposure, many individuals—with varying levels of exposure to a CBRNE agent—experience nonspecific (and also specific health-related) stress-related symptoms, most of which will subside in the short term. It is also important to recognize that individual differences dominate event characteristics in the response to toxic events and treatment options [21]. Hence, new tools must be developed, which will largely be based on objective agent-specific symptoms [22]. Nonetheless, elements of various psychological assessments might prove useful in community crisis awareness tools to reflect the level of distress by the individual and assist in the differential diagnosis.

Thus, automated crisis health risk self-assessment tools will primarily aim to discriminate between agent-derived illness and psychogenic phenomena through: 1) knowledge of the range of symptomatology generated by given biothreat agents over time, 2) our expanding knowledge and database of sociodemographic and psychosocial factors that characterize persons demonstrating MUPS [23]–[26], and 3) geospatial and temporal data of actual threat exposure information. Further sensitivity and specificity will be achieved by including questions to assess individual response proclivities, including nonattentive or social desirability answering, minimizing, exaggerating, and frank malingering. The result will be a probabilistic differentiation between those who are suffering from actual exposure and those without clear exposure but who are nonetheless affected by an event. With time and validation testing, such systems could become further sophisticated and capable, but even an imperfect system would be an enormous improvement to the current system, or lack thereof, in which no attempt is made to segregate the MUPS from exposed individuals.

An ideal assessment tool would be tiered, sensitive to time and event processes, and interactive. As an example of a tiered assessment, someone who self-reports symptoms of confusion would move down the algorithm to a brief assessment of anxiety, which could exacerbate these symptoms. If such symptoms are not anxiety related, the person could be given a brief automated neurocognitive screening battery.

Such tools could not only categorize symptoms but also assess the severity and phase of anxiety or illness. Interactive discernment of levels of possible exposure would lead to various agent-specific algorithms in tiered systems. This information would be useful for prioritization schema, as well as useful for specific medical care or mental support. Thus, a severity scale or time plot wizard would assist emergency or mental healthcare workers to address patient needs optimally and efficiently, as well as to furnish a picture of overall community health.

Not only might this novel technology prove to be more efficient, but it might also improve the accuracy of discrimination and differential diagnosis. Emergency physicians are not generally adept at dealing with psychogenic symptomatology [27], whereas the computer can deploy sophisticated heuristics and statistics to analyze the responses of a presenting individual, and use dynamic real-time state-of-the-art information that has evolved during a particular biological event.

TABLE 1. Examples of existing and possibly relevant psychometrics.*

Generalized Psychosocial Distress:

- Symptom-Checklist-90 (SCL-90) (Derogatis, 1973)
- Brief Symptom Inventory (BSI) (Derogatis and Melisaratos, 1983)
- Primary Care Evaluation of Mental Disorders (PRIME-MD) (Spitzer et al., 1994)
- General Health Questionnaire (GHQ) (Goldberg, 1972)
- Symptom-Driven Diagnostic System Primary Care (SDDS-PC) (Broadhead, 1995)
- Mental Health Inventory (MHI) (Veit and Ware, 1983)
- Behaviour and Symptom Identification Index (BASIS-32) (Eisen and Grob, 1989)
- Impact of Events Scale (IES) (Horowitz, Wilner, and Alvarez, 1979)

Depression:

- Beck Depression Inventory (BDI) (Beck et al., 1961)
- Center for Epidemiological Studies (CES-D) (Radloff et al., 1977)
- Hamilton Rating Scale for Depression (HAM-D) (Hamilton et al., 1967)

Post Traumatic Stress Disorder (PTSD):

- PTSD Checklist (PCL) (Weathers et al., 1991)
- PTSD Symptom Severity Interview (PSSI) (Foa et al., 1993)
- Clinician-Administered PTSD Scale (CAPS) (Blake et al., 1990)

Illness Worry and Conviction:

- Whiteley Index (Pilowsky, 1967)
- Illness Attitudes Scale (Kellner, 1986)
- Penn State Worry Questionnaire (PSWQ) (Meyer, Miller, Metzger, and Borkovec, 1990)

Stress:

- Perceived Stress Scale (PSS) (Cohen, 1983)

Anxiety:

- Anxiety Screening Questionnaire (ASQ) (Wittchen and Boyer, 1998)
- Hamilton Anxiety Scale (HAS) (Hamilton, 1959)

** These are provided as examples; there are numerous other validated and accepted instruments.*

Current Efforts

We now describe three complementary existing or developing tools that aid or could aid individuals in conducting self-assessments following a bioterrorist outbreak.

Proposed National Capital Region Self-Risk Assessment and Communication Demo System

This proposed self-risk assessment and communication system is based upon an initial assessment tool developed at the Harvard School of Public Health and the School of Medicine, by Rick Newell, Steven Locke, Paul Testa, and Kathy Wong. The plan is to launch the system on Global Health Initiatives servers as a testbed within the National Capital Region during 2004 [28]. The function of this tool is to intervene as a triage mechanism in the period following a likely biological weapon attack. Secondarily, it assists with the separation of those who are manifesting symptoms from actual exposure to the agent from those who are experiencing the symptoms as a result of attack-related stressors. Risk communication skills are used throughout the system to appropriately convey the relative risk of exposure. An incorporated geographic information system (GIS) allows the tool to assist with public health outbreak investigations, case finding, and directions customizable to the individual taking the assessment. The tool can be deployed via paper, phone with speech-recognition software, home computer via the Internet, a screening center, or an acute care facility. The assessment begins with the questions most predictive of actual exposure. If the respondent answers these questions in such a way that indicates he or she should seek medical care, the program will so direct the individual, and the assessment will end.

This tool assesses four types of data. Type 1 data assesses the patient demographics and situational factors such as proximity to exposure and availability of mass transportation. Type 2 data are customized to a given attack/agent, based upon knowledge of the presumptive agent from syndromic surveillance, intelligence, or detection systems. Type 3 data assess symptoms that are more likely the result of acute stress rather than exposure. Type 4 data assess the susceptibility of the individual to experience physical symptoms as a result of stress. A final score is given to each patient and reflects the probability that he or she is experiencing symptoms related to exposure versus the acute stress of the attack. A threshold score will be set at which the respondent will be directed to seek medical care in order to be treated for exposure or further rule out exposure. Other thresholds can be set directing respondents to other sources of care or counseling. These thresholds will be adjusted during the time following an attack based on the capacity and flow through emergency and healthcare services. The threshold can be lowered continuously as healthcare services are able to deal with added patients, and thus all respondents can be seen eventually if they so desire. By delaying or smoothing out the somatization surge post event, medical access of those more likely at risk is protected and ensured.

Pediatric Disaster Systems of Care and PsySTART

PsySTART is a pediatric disaster system of care model linked to a technology-based rapid triage system [29], [30]. Developed by Merritt Schreiber at the David Geffen School of Medicine at UCLA, the system is designed for comprehensive incident management of pediatric mental health risk across phases of mass casualty events and across various "disaster systems of care." The system uses a Web-based interface to systematically triage, assess, and track children at risk from psychological consequences of mass casualty events, disasters, and weapons of mass destruction (WMD) terrorism events. PsySTART provides real-time, GIS-enabled linkage between emergency departments, primary care providers, schools, specialized medical disaster settings (i.e., NDMS /DMAT), disaster relief agencies, and public mental health to link services for children at high risk for post-event effects. It serves as a centralized communication tool to coordinate the acute phase and recovery mental health response and establishes a common protocol for seamless triage, screening, clinical assessment, and definitive care.

Automated Telephone-Based Symptom Triage

TLC is a telecommunications system for monitoring patient symptoms and other clinical findings and for educating, advising, and counseling patients and consumers about their health and medical conditions. TLC technology and clinical applications have been developed and evaluated by Robert Friedman and colleagues at Boston University with about $20M of support from the National Institutes of Health and the Agency for Healthcare Research and Quality over the past 20 years. TLC systems use validated instruments supplemented by expert clinician input to evaluate potential symptoms and to report these automatically to responsible clinicians in a form that assists in clinical decision making [31]–[34]. For a future biodefense preparedness system, TLC could be used, via a toll-free telephone number, for people who believe they may have been affected in an event to triage them for appropriate follow-up. Using behavior change intervention strategies that have been effective in a variety of health behaviors, TLC could also educate, advise, and counsel individuals who are likely to have stress-related symptoms.

System Architecture

Automated health risk self-assessment tools should be incorporated into IT system software that is sufficiently general so it can be applied to jurisdictions throughout the United States, and yet is sufficiently flexible to be customizable for local needs. Ultimately, any IT solution to bioterrorist incidents should be a distributed system, based upon core standardized and interoperable tools that allow customization and localization at the state and local levels. Any IT software must be flexible to accommodate to new, dynamic, and fluctuating situations, as well as to adapt to a variety of jurisdictional political requirements and information system architectures. Open-system software should be used to permit the greatest flexibility for adaptation. The system must have a central architecture with the core elements of the system, as well as local controllables, supplemental databases, and customized applications that can be added over time. The system should be redundant so that it is not dependent upon only one communication mode; for instance, it would not be wise to assume the Internet will be available for the system to function during a crisis.

Acceptance and integration into the IT fabric of modern governments will require consideration of pertinent standards. The broader biodefense and homeland security information infrastructure to which these system components may be linked will also play key roles in assisting outbreak responses. The system may integrate multiple functional components (e.g., self-assessment, resource tracking, referral, risk communication, surveillance, alerts, geographic tracking) into a seamless decision support system. Thus, a broad variety of applications—such as computer-assisted screening interviews, meme [35] surveillance of portal databases to trigger alerts on psychological and behavioral issues, chemical and biological

The numbers of individuals who seek care in the absence of exposure may be orders of magnitude larger than the number of individuals seeking care who are actually exposed or ill.

detection devices, monitoring contact distance through the use of locator bracelets, tracking of patients, allocation of resource stockpiles, management of patient flow, and tracking of behavioral response to risk communications—can and should be developed and linked to the self-risk assessment and communication IT system. (A "meme" is an idea that is passed on from one human generation to another. It is the cultural equivalent of a gene, the basic element of biological inheritance. The term was coined in 1976 by Richard Dawkins in his book, *The Selfish Gene*. Dawkins speculated that human beings have an adaptive mechanism that other species don't have. In addition to genetic inheritance, with its possibilities and limitations, humans, said Dawkins, can pass their ideas from one generation to the next, allowing them to surmount challenges more flexibly and more quickly than through the longer process of genetic adaptation and selection.) In this way, multifunctional homeland security system components can interconnect and work synergistically to provide a seamless fabric that best protects our citizens and societal infrastructure.

We propose that the system infrastructure conform to the Department of Homeland Security (DHS) Federal Enterprise Architecture (FEA) Technical Reference Model Version 1.0 [36] and be informed by Department of Defense biosurveillance data integration efforts. In addition, the software should be compatible with the developing National Health Information Infrastructure [37] following the specific guidance for IT systems [38] and risk communication [39] established by the Centers for Disease Control and Prevention (CDC). It is anticipated that the architecture may rely heavily on messaging infrastructures (i.e., message queuing) and content (XML, HL7, etc.) while allowing for fully autonomous local functioning and capturing clinical data using the expected standards (e.g., LOINC, SNOMED, and DICOM.

Security and privacy concerns will be of utmost importance to any such system that deals with sensitive individual information and public surveillance data. Universal secure access to the system is necessary, however, for participation of government leaders, responders, and the public.

Other System Functions

In addition to the primary purpose of health risk self-assessment screening and reduction of surge, there are other significant potential functions of this proposed system: 1) medical staging and prioritization, 2) contact distance management, 3) surveillance of community mental health and psychosocial/behavioral concerns, and 4) risk communication.

Medical Staging and Prioritization

In addition to discriminating psychogenic manifestations from bioagent-derived symptoms, the system could also provide information on the level of emotional distress, a likely phase of the malady that would be useful for later emergency caregivers to have at their fingertips. Self-assessment tools could provide a tiered structure of screenings according to constructed algorithms that would order clinical diagnostic tests automatically on the basis of respondents' answers.

Contact Distance Management

The crisis self-assessment tools may provide the epidemiological data necessary for public health interventions to prevent contact with an infectious agent and to maintain social distance between exposed and nonexposed groups. In the case of communicable infectious disease, the benefit is clear and compelling. In addition, it is also good medical practice to separate those with actual disease from those with stress disorders and other psychopathologies to prevent further psychogenic ("hysterical") contagion. An automated crisis health risk self-assessment system could play a major role in increasing the efficiency of patient isolation and self-quarantine in the case of community spread of an infectious agent such as severe acute respiratory syndrome (SARS).

Risk Communication

The proposed automated crisis health risk self-assessment system can be used as a vehicle to facilitate communications to the general public. The quality of the communication about an outbreak and corresponding messages of civic leaders and government officials can foster fear or allay anxiety, and thus the resultant levels of stress, distress, and MUPS [40], [41]. For example, Winston Churchill through his oratory effectively calmed the British people in the face of great danger. The ability to identify, categorize, prioritize, and communicate risks is critical to optimize response in disasters [42].

The proposed system could be a significant component of a wider community risk communication effort. The communicators used in automated crisis health risk self-assessment communication components should be known and trusted figures within a given community (e.g., video sequencing showing a respected person advising you to stay calm or helping you consider whether you have symptoms pertinent to the current emergency). The IT system could educate and inform specific targeted communities by providing linkages to other local sources of information. It would help augment understanding of basic biological concepts such as relative risk, exposure, contagion, and so forth, so that individuals may make better and more empowered decisions regarding their health in crises. The interactive nature of this technology could be extended into providing mini-interventions for various problems (e.g., lack of information or understanding about a biological agent) and individually tailored on-the-spot corrections of misinformation, reducing uncertainty, and decreasing anxiety.

Automated components of these systems will enable automatic translation of messages into other languages and use culturally competent and educationally appropriate vocabulary and phraseology. IT permits highly interactive and customizable communications that enable appropriate responses to an individual's and subpopulation's reactions in helping make health- and life-critical decisions and take actions in high-risk, and potentially low-trust, environments. A person could indicate in the beginning of the assessment whether s/he would like to have the "interviewer's" or communicator's voice in a certain language or gender, and the system could recognize the source location and customize accordingly. Groups that might receive specifically tailored messages include those immediately threatened by the incident, healthcare workers, rural dwellers, families, and those with special needs. Such a system could deliver evidence-based messages that are soothing as well as action oriented, precisely to those most concerned. To address the individual's need for self-preserving actions in a context of understanding issues affecting the common good, programs can be designed to deliver messages designed to maintain and improve community cohesion and resilience in times of emergency and crisis.

Surveillance of Community Mental Health

Simultaneously, while the system assists preliminary screening functions, data can be generated on the evolving public reaction to a critical incident. Key professional, community, and government decision makers will be able to monitor surge capacity, public response, and message structure in a secure back-end component of the system that includes scientific visualization and decision support employing a GIS. Information on the number of patients directed to mental health workers, emergency medical care, and other support systems will facilitate ideal healthcare resource allocation. Moreover, the data will be a real-time direct gauge of the true incidence and magnitude of public concern and psychosocial dimensions, along with geospatial and temporal aspects used to determine trending. Viewing this information dynamically on GISs enables key decision makers to respond effectively to the most pressing considerations of the crisis at that specific moment. Specifically, the system could permit a more formalized community-wide risk assessment to public health and emergency services. The need for such a system was highlighted during the 2002 anthrax letter attacks, when the CDC received over 11,000 telephone calls from the public. The absence of standardized health communication protocols prevented efficient follow-up investigations and could not give an up-to-the-minute sophisticated picture of the public reaction [43].

Implementation

Stakeholders

Automated health risk self-assessment software components are likely to be developed by a broad array of public and private interests. A range of licensing and ownership options are likely to emerge if automated health risk self-assessment toolkits and systems are determined by pure market forces. System service stakeholders might include states, political jurisdictions, public health departments, mental health services, emergency services, nongovernmental organizations (e.g., the American Red Cross), regional and local hospitals and healthcare systems, as well as the patients themselves. It is suggested that core system components be developed according to national standards, evaluated by the Federal government, but maintained by private interests required to meet established performance and safety standards. Incentives, such as from an advanced component economy, should be created to ensure that automated health risk self-assessment tools, components, and systems continue to evolve in depth, breadth, and quality. Attempts should be made to benchmark the most effective self-assessment tools, components, and systems.

Emergency services have command and control authority over crisis incidents and thus they are a possible service owner, particularly if the system is aimed at prehospital distribution. The data generation and risk communications aspects of such a system would seem clearly to fit within governmental crisis response efforts. Medical triage, however, is an important entry function of healthcare systems, and automated health risk self-assessment tools could play a part in directing individuals toward appropriate hospital admitting operations. Healthcare systems will want control over patients once they are entered into their system. A psychosocial or mental health triage tool may be claimed as the province of human resource departments. Perhaps the most obvious owner/licensee would be public health departments, in that community health and the medical consequences of a bioterrorist outbreak are usually the province of public health authorities. It is conceivable, however, that public health systems may not see such an IT solution as among their traditional duties and responsibilities. We would argue that the need to prevent the overload of medical services of a region—which could lead to a failure of system response—is logically construed as a local or regional public health responsibility that transcends any single healthcare system.

Regardless of the agency that takes ownership of the system, it would be helpful for the effort to be communicated to and reviewed by as many stakeholders as possible, including the medical community, public health, emergency responders, and members of the community whose suggestions would improve acceptability and feasibility of implementation. Challenges for obtaining cooperation will include funding, "turf," and the difficulty in validating the early instantiations of the systems.

Liability

Liabilities are created both from errors of commission and errors of omission. Some may voice concerns over the liability involved in self-assessment tools. This same concern was voiced in earlier years with respect to medical Web sites and self-help books, which have now been in common usage for many years without significant problems regarding litigation. Nonetheless, national standards establishing the scientific basis and validation and verification process for such systems will be important to maximize public safety and minimize liability for the product developers and service owners/licensees who comply with the national standards.

The preliminary assessment, screening, referral, and risk communication components should carry disclaimers that the system is for educational purposes and is not to be considered diagnostic (which would, under the present conditions, require the direct intervention of a physician in real time). However, in lieu of full diagnosis, there is reason to believe from the extensive literature on premedical screening that the assessment and

referral components can handle large numbers of people in a premedical screening mode that directs those with high risks of exposure or medical symptoms to appropriate medical and mental health services for further diagnosis and treatment.

Upon completion of the screening, the symptomatic patient will end up in the care of a traditional human service-based medical or mental health professional. If the individual has been found to be misdirected, he or she can be then be redirected to the more appropriate medical element. The greatest concern is that individuals who are at high risk might use health risk self-assessment tools and not seek necessary care. This potential problem needs to be carefully studied in controlled environments—perhaps in nonemergency conditions, such as during the annual flu season—to explore the extent of this potential problem. Every effort should be made to avoid false negative risk designations while also understanding the critical need to reduce infrastructure overload by MUPS patients and lower-risk, asymptomatic individuals.

Without such a premedical screening system and with medical systems under the stress of infrastructure overload, many people in need of immediate medical care may be delayed in reaching medical care or may not reach a medical professional during the critical time necessary to optimize treatment and reduce harm to others. Before community crisis assessment systems come into common usage, the perceived liability will be in the adoption of such systems. Once systems are in common practice, by far the greater liability will be in *not* providing access to health risk self-assessment tools.

Pilot Projects

Pilot testing will be necessary for validation and verification efforts to garner acceptance and to make improvements in the self-assessment tools and IT system. With pilot tests, many problems, gaps, and potential solutions will be revealed. Although the system's objectives are compelling, early testbeds must be carefully evaluated and well documented in attaining their proposed goals before many jurisdictions will adopt the operational systems. Validation will, in the first instance, begin with clinical practice and perhaps also during emergency preparedness training exercises. Simulated attacks have been used to assess community reactions. For example, DiGiovanni and colleagues simulated an intentional Rift Valley fever outbreak in a community in the southern part of the United States using a series of simulated print and television "news reports" over a fictional nine-day crisis period [44].

Demonstrations through early adopters will be necessary. Potential testbeds might include: 1) the National Capital Region; 2) Allegheny County, Pennsylvania; 3) Santa Clara County, California; 4) Los Angeles County, California; and 5) the state of Massachusetts, with a focus on the Boston metropolitan area. In each area, there is a very active bioterrorism preparedness effort, perceived vulnerabilities to infrastructure overload, and a willingness to investigate new IT solutions.

Funding

DHS may also be a significant source of funding for prehospital systems to states and localities. CDC and the Health Resources and Services Administration funds may also be useful in distributing toolsets as regional patient surge capacity/bioterrorism equipment. The need for such systems would seem to be a national mandate, in the wake of 911 and the anthrax outbreak of fall 2001. The cost of effective risk assessment tools and communication systems amortized over many jurisdictions would be a small fraction of the human service costs associated with addressing the service need directly through the health care system, let alone the human costs and economic damage associated with a healthcare system failure during a crisis. Thus, the need to proliferate and distribute effective automated health risk self-assessment systems broadly should follow quickly.

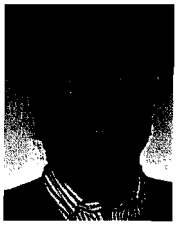
Conclusion

To date, biological weapons have proven to be more of a psychological threat than a physical danger, and while they may someday result in significant mortality they would seem always to create larger numbers of psychological illness. At least in the United States, they have proven to be weapons not of mass destruction but of mass psychogenic illness. Hippocrates first noted this phenomenon around 400 BC when he introduced the term *hysteria* to describe a group of women he observed who had unexplained muscle spasms, abdominal cramps, nausea, and headaches [45]. This phenomenon has been recognized as the presentation of physical symptoms for which there is no underlying physical cause but instead appear to be manifestations of psychological distress. Healthy adults commonly experience isolated minor symptoms, minor illness, and transient dysfunctions in the course of daily life that could be construed as within the range of hysterical phenomena. Population-based surveys indicate that 86% to 95% of the general population experience at least one symptom in a given two- to four-week interval, and the typical adult has at least one somatic symptom every four to six days [46]–[50]. After a bioterrorist attack, individuals experiencing these psychogenic symptoms will outnumber those actually exposed by many times. These two patient groups must be separated and addressed differently; yet, there is currently no established, validated tool or mechanism to manage this realistic challenge. By using the power of information technology, self-assessment tools will be able to distinguish these two groups; provide initial customized feedback to individuals about their risk, what they can do to manage stress and psychogenic symptoms, and/or seek medical care for likely agent or event-based physical symptoms; provide a mechanism to screen, prioritize and distribute the patient surge; optimize resource allocation; facilitate individual and community risk communications; and surveil and monitor public exposure, stress, and reactions to crisis events.

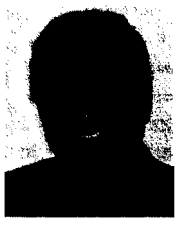
Victor W. Weedn is a member of the IEEE Bioterrorism and Genomics and Bioinformatics Working Groups. He is a principal research scientist and the director of Biotechnology and Health Initiatives at Carnegie Mellon University. He is best known for creation of the U.S. military's DNA identification program. In this position, he helped shepherd early DNA microchip technology within the military and thus became involved in bioagent detection technologies. He is a former codirector of the CMU/UPitt BioMedical Security Institute. Dr. Weedn received his B.S. in psychology at the University of Texas at Austin; his M.S. at the UT Southwestern Medical School in Dallas; and completed his pathology residency at Baylor College of Medicine and the UT Houston Medical School, his

anatomic pathology fellowship at the MD Anderson Hospital and Tumor Institute, and his forensic pathology fellowship at the Armed Forces Institute of Pathology.

Michael D. McDonald is a chairman of the Genomics and Bioinformatics Working Group, a member of the Bioterrorism Working Group and Medical Technology Policy Committees of the IEEE, and has coordinated several strategic planning and legislative education efforts in the area of biodefense communications and computing infrastructure. He is the coordinator, Health and Biodefense, of the Georgetown Life Sciences and Society Initiative and chairman, Communications Committee National Capital Region, Emergency Response/EF. He is CEO and president of Global Health Initiatives, Inc., founding chairman of HealthCentral.com, and past managing director of Health for the Koop Foundation. Over the past 15 years, Dr. McDonald has developed several computer-based tools in the area of health risk assessment. These health risk assessments have now been used by millions of people to assess their health and the health of their institutions. Dr. McDonald is a graduate of the School of Public Health of the University of California Berkeley.



Steven E. Locke is the founder and president of Veritas Health solutions, a Boston-based behavioral medicine and disease management consulting firm. Dr. Locke teaches at Harvard Medical School as an associate professor of Psychiatry and at Massachusetts Institute of Technology, where he is an associate professor of Health Sciences and Technology. He most recently was the chief of Behavioral Medicine at Harvard Vanguard Medical Associates, where he led the development of disease and demand management programs. Dr. Locke received his medical degree from Columbia University College of Physicians and Surgeons and completed his residency in psychiatry at Harvard with additional training at Boston University School of Medicine. He is the author or editor of five books and 37 book chapters and journal articles. He is currently president of the American Psychosomatic Society, chair of the Technology Committee of the Disease Management Association of America, and chair of the Massachusetts Psychiatric Association's Task Force on Primary Care. Since September 11, 2001, he has become involved in public health preparedness related to psychosocial aspects of disasters, terrorism, and outbreaks. In 2002, Dr. Locke was invited to join the Massachusetts Department of Public Health's program on bioterrorism preparedness.



Merritt Schreiber is a child psychologist in the David Geffen School of Medicine at UCLA. He serves as the program manager for the Terrorism/Disaster Branch of the National Center for Child Traumatic Stress co-located at UCLA and Duke. He received his Ph.D. in clinical child psychology from Ohio State University in 1982 and was a postdoctoral Fellow in the University of Southern California School of Medicine, Department of Psychiatry. Since then, he has served in many capacities and now focuses on coping behaviors in children exposed to disaster, including treatment models, clinical data management, mental health, and clinical comorbidity. He was appointed by U.S. Secretary Tommy Thompson to the Emergency Public Information and Communications Advisory Board under the Bioterrorism Public Health Preparedness and Security Act of 2002. He received an APA presidential citation for his work with families who lost loved ones on September 11. He is Southern California coordinator, Disaster Response Network, California Psychological Association and on the American Red Cross National Critical Response Team.

Robert H. Friedman is the chief of the Medical Information Systems Unit at Boston Medical Center and professor of medicine and public health at Boston University Schools of Medicine and Public Health. Dr. Friedman completed his medical education at Stanford University. His research group has pioneered the application of information science to monitoring and changing important health behaviors: eating, exercise, cigarette smoking, medication taking, and monitoring patients with chronic disease and promoting their self-care. The laboratory developed a unique telecommunications technology, Telephone-Linked Care (TLC), that carries out totally automated, computer-controlled telephone conversations with patients in their homes. This TLC technology has been used to assess symptoms, function, and health status of patients and consumers and to modify their health behavior and thus improve health outcomes.

Richard G. Newell is completing his masters in public health at Harvard University. His experience includes the design of a Web-based triage tool to assist with the separation of persons experiencing symptoms due to exposure to a bioterrorist agent from those experiencing symptoms due to the acute stress of the situation, and to assist public health officials with the outbreak investigation. While at Harvard, he also held responsibilities as a teaching fellow and a health care management intern at Harvard University Health Services. Additionally, he has also been employed by Chautauqua County Health Department, where he worked with both communicable disease reporting and county and state environmental health code enforcement. He is a co-founder of the Lighthouse Free Medical Clinic, located on Buffalo's east side. He is completing his fourth year of medical education at the State University of New York at Buffalo. Rick will be commencing an emergency medicine residency in 2004, where he will continue his work in disaster medicine.

Lydia R. Temoshok directs the Behavioral Medicine Program at the Institute of Human Virology (IHV) and is a professor of medicine and psychiatry at the University of Maryland School of Medicine at Baltimore. Before coming to the IHV in 1998, Dr. Temoshok served as senior scientist with the World Health Organization in Geneva, Switzerland, and directed HIV

behavioral research for the U.S. military. While on the faculty at the University of California School of Medicine in San Francisco (1978–1989), she directed biopsychosocial research on the role of emotion and coping in the progression of malignant melanoma, and in 1982 formed and led a multidisciplinary group of scientists to study a new syndrome that later would be called AIDS. Dr. Temoshok has published ten books or monographs, over 40 book chapters, and over 100 journal articles in the areas of biopsychosocial oncology, psychoneuroimmunology, and behavioral factors relevant to HIV prevention.

Address for Correspondence: Victor W. Weedn, Carnegie Mellon University, 4400 Fifth Ave., Pittsburgh, PA 15213 USA.

References

[1] R.E. Bartholomew and S. Wessely, "Protean nature of mass sociogenic illness: From possessed nuns to chemical and biological terrorism fears," *Br. J. Psychiat.*, vol. 180, pp. 300–306, 2002.
[2] C.C. Engel, "Outbreaks of medically unexplained physical symptoms after military action, terrorist threat, or technological disaster," *Mil. Med.*, vol. 166, suppl. 2, pp. 47–48, 2001.
[3] R.H. Gurwitch, M. Kees, S. Becker, M. Schreiber, B. Pfefferbaum, and D. Diamond, "Understanding and responding to the mental health needs of children following disasters," *J. Prehospital and Disaster Med.*, in press, 2004.
[4] T.F. Jones, A.S. Craig, D. Hoy, E.W. Gunter, D.L. Ashley, D.B. Barr, J.W. Brock, and W. Schaffner, "Mass psychogenic illness attributed to toxic exposure at a high school," *New Eng. J. Med.*, vol. 342, pp. 96–100, 2000.
[5] N. Asukai and K. Maekawa, "Psychological and physical health effects of the 1995 sarin attack in the Tokyo subway system," in *Toxic Turmoil. Psychological and Societal Consequences of Ecological Disasters*, J.M. Havenaar, J.G. Cwikel, and E.J. Bromet, Eds. New York: Kluwer Academic/Plenum Publishers, 2002. pp. 149–162.
[6] J.S. Petterson, "Perception vs. reality of radiological impact: The Goiania model," *Nuclear News*, vol. 31, pp. 84–90, 1998.
[7] F.M. Burkle, "Acute-phase mental health consequences of disasters: Implications for triage and emergency medical services," *Ann. Emerg. Med.*, vol. 25, pp. 119–128, 1996.
[8] P.R. Jenkins, R.A. Jenkins, E.D. Nannis, L.T. McKee, and L.R. Temoshok, "Reducing sexually transmitted disease (STD) and HIV risk in a military STD clinic: Evaluation of a randomized preventive intervention trial," *Clin. Infect. Dis.*, vol. 30, pp. 730–735, 2000.
[9] M. Morris, D. Sparrow, D. Gottlieb, and A. Selim, "A telephone linked computer system for COPD care," *Cancer Practice*, vol. 10, no. 3, pp. 147–154, 2001.
[10] H.K. Delichatsios, R.H. Friedman, K. Glanz, S. Tennstedt, C. Smigelski, B.M. Pinto, H. Kelley, and M.W. Gillman, "Randomized trial of a 'Talking Computer' to improve adults' eating habits," *Amer. J. Health Promotion*, vol. 15, no. 4, pp. 215–224, 2001.
[11] K.L. Jarvis, R.H. Friedman, T. Hereen, and P.M. Cullinane, "Older women and physical activity: Using the telephone to walk," *Women's Health Issues*, vol. 7, no. 1, pp. 24–29, 1997.
[12] R.H. Friedman, L.E. Kazis, A. Jette, M.B. Smith, J. Stollerman, J. Torgerson, and K.A. Carey, "A telecommunications system for monitoring and counseling patients with hypertension. Impact on medication adherence and blood pressure control," *Am. J. Hypertension*, vol. 9. no. 4, pt. 1, pp. 285–292, 1996.
[13] B.M. Pinto, R. Friedman, B.H. Marcus, H. Kelley, S. Tennstedt, and M.W. Gillman, "Effects of a computer-based telephone counseling system on physical activity," *Am. J. Preventive Medicine*, vol. 23, pp. 113–120, 2002.
[14] D.F. Mahoney, B.J. Tarlow, and R.N. Jones, "Effects of an automated telephone support system on caregiver burden and anxiety: Findings from the REACH for TLC intervention study," *Gerontologist*, vol. 43, no. 4, pp. 556–567, 2003.
[15] J. Briere, "*Psychological assessment of adult posttraumatic states*," Washington, D.C.: American Psychological Assoc., 1997.
[16] R.L. Kane and G.G. Kay, "Computerized assessment in neuropsychology: A review of tests and test batteries," *Neuropsychol. Rev.*, vol. 3, pp. 1–117, 1992.
[17] M. Horowitz, M. Wilner, and W. Alvarez, "Impact of event scale: A measure of subjective stress," *Psychosomatic Medicine*, vol. 41, pp. 209–218, 1979.
[18] D. Weiss and C. Marmar, "The impact of event scale–Revised," in *Assessing Psychological Trauma and PTSD*, J. Wilson and T. Keane, Eds. New York: Guildford, 1997.
[19] G.J. Devilly, "Assessment devices," 1997. Available: http://www.criminology.unimelb.edu.au/victims/resources/assessment/ assessment.html
[20] C.R. Brewin, "Screening for posttraumatic Stress Disorder after exposure to trauma," review and discussion paper for NIMH/SAMHSA/NCPTSD Roundtable Meeting on Assessment and Screening Following Disasters and Mass Violence, Aug. 26–27, 2003.
[21] M.L. Bowman, "Individual differences in posttraumatic distress: Problems with the DSM-IV model," *Can. J. Psychiatry*, vol. 44, no. 1, pp. 21–33, 1999.
[22] L. Temoshok, "Proposed strategies and technologies to differentiate biological and psychogenic reactions to bioweapons, and to detect early neurocognitive and behavioral sequellae," draft report available from Dr. Temoshok, Institute of Human Virology, 725 West Lombard St., Baltimore, MD. 21201; temoshok@umbi.umd.edu.
[23] L.S. Zun, "Panic disorder: Diagnostic and treatment in emergency medicine," *Ann. Emerg. Med.*, vol. 30, no. 1, pp. 92–96, 1997.
[24] "Psychological impact of recent terrorist attacks." Available: http://www.facingfeartogether.org/healthtogether/facingFear/litReview.html
[25] "Somatoform disorders," in *The Merck Manual of Diagnosis and Therapy*, 17th ed. Whitehouse Station, NJ: Merck & Co., Inc., 2003.
[26] *Diagnostic and Statistical Manual of Mental Disorders* (DSM IV). Washington, DC: American Psychiatric Assoc., 1994.
[27] D. Goldberg, "Reasons for misdiagnosis," *Psychological Disorders in General Medical Settings*, N. Sartorius, D. Goldberg, G. de Girolamo, J. Costa e Silva, Y. Lecrubier, and U. Wittchen, Eds. Toronto: Hogrefe & Huber, 1990.
[28] M. McDonald, personal communication, Georgetown Working Group, June 2003.
[29] M.D. Schreiber, "Children's emergencies in weapons of mass destruction and terrorism: Disaster system of care, rapid triage, and consequence management," paper presented at the *16th Annu. California Injury Prevention Conf.*, Sacramento, CA, 2002.
[30] R.S. Pynoos, A.M. Steinberg, M.D. Schreiber, and M.J. Brymer, "Children and families: A new framework for preparedness and response to danger, terrorism and trauma," in *Group Approaches for the Psychological Effects of Terrorist Disasters*, in press, 2004
[31] R.H. Friedman, "Automated telephone conversations to assess health behavior and deliver behavioral interventions," *J. Medical Syst.*, vol. 22, no. 2, pp. 95–102. 1998.
[32] R.H. Friedman, J. Stollerman, L. Rozenblyum, D. Belfer, A. Selim, D. Mahoney, and S. Steinbach, "A telecommunications system to manage patients with chronic disease," *Medinfo*, vol. 9, pt. 2, pp. 1330–1334, 1998.
[33] A.C. King, R.H. Friedman, B. Marcus, C. Castro, L. Forsyth, M. Napolitano, and B. Pinto, "Harnessing motivational forces in the promotion of physical activity: The community health advice by telephone (CHAT) project," *Health Educ. Res.*, vol. 17, no. 5, pp. 627–636, 2002.
[34] H.R. Ramelson, R.H. Friedman, and J.K. Ockene, "An automated telephone-based smoking cessation education and counseling system," *Patient Education and Counseling*, vol. 36, pp. 131–144, 1999.
[35] Available: http://whatis.techtarget.com/definition/0,sid9_ gci212545,00.html
[36] Federal Enterprise Architecture (FEA) technical reference model version 1.0. Available: http://www.feapmo.gov/fea_budget_formulation_docs.asp
[37] *NHII—Information for Health: A Strategy for Building the National Health Information Infrastructure*. Washington, D.C.: National Committee on Vital and Health Statistics, Nov. 15, 2001. Available: http://ncvhs.hhs.gov/reptrecs.htm
[38] Available: http://www.bt.cdc.gov/planning/continuationguidance/pdf/appendix-4.pdf
[39] Available: http://www.bt.cdc.gov/planning/continuationguidance/pdf/communication_ attachf.pdf
[40] H.C. Holloway, A.E. Norwood, C.S. Fullerton, C.C. Engel, and R.J. Ursano, "The threat of biological weapons: Prophylaxis and mitigation of psychological and social consequences," *JAMA*, vol. 278, pp. 425–427, 1997.
[41] G.M. Gray and D.P. Ropeik, "Dealing with the dangers of fear: The role of risk communication," *Health Aff.*, vol. 21, no. 6, pp. 106–116, 2002.
[42] C.N. Haas, "The role of risk analysis in understanding bioterrorism," *Risk Anal*, vol. 22, no. 4, pp. 671–677, 2002.
[43] J.A. Mott, T.A. Treadwell, T.W. Hennessy, P.A. Rosenberg, M.I. Wolfe, C.M. Brown, and J.C. Butler, "Call-tracking data and the public health response to bioterrorism-related anthrax," *Emerg Infect Dis.*, vol. 8, no. 10, pp. 1088–1092, 2002.
[44] C. DiGiovanni, Jr., B. Reynolds, R. Harwell, E.B. Stonecipher, and F.M. Burkle, Jr., "Community reaction to bioterrorism: Prospective study of simulated outbreak," *Emerg. Infect. Dis.*, vol. 9, no. 6, pp. 708–712, 2003.
[45] L. Temoshok and C.C. Attkisson, "Hysterical phenomena: Epidemiology and evidence for psychosocial theory," in *The Hysterical Personality*, M.J. Horowitz, Ed. New York: Jason Aronson, 1997, pp. 143–222.
[46] K. Dunnell, "Medicine takers and hoarders," *J. Royal College of General Practitioners*, vol. 23, no. 2, pp. 2–9, 1973.
[47] M. Wadsworth, "Health and sickness: The choice of treatment," *J. Psychosomatic Res.*, vol. 18, no. 4, pp. 271–276, 1974.
[48] K.L. White, T.F. Williams, and B.G. Greenberg, "The ecology of medical care," 1961. *Bull. New York Acad. Medicine*, vol. 73, no. 1, pp. 187–205, 1996. *New Engl. J. Med.*, vol. 265, no. 18, pp. 885-892, 1961.
[49] K. Dunnell, "Medicine takers and hoarders," *J. Royal College of General Practitioners*, vol. 23, no. 2, pp. 2–9, 1973.
[50] K.J. Egan and R. Beaton, "Response to symptoms in healthy, low utilizers of the health care system," *J. Psychosomatic Res.*, vol. 31, no. 1, pp. 11–21, 1987.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Drug and Alcohol Dependence 76S (2004) S21–S29
DRUG and ALCOHOL DEPENDENCE
www.elsevier.com/locate/drugalcdep

Suicidal behavior, drug use and depressive symptoms after detoxification: a 2-year prospective study☆

James D. Wines Jr. (MD)[a,*], Richard Saitz (MD, MPH)[b,c], Nicholas J. Horton (ScD)[d], Christine Lloyd-Travaglini (MPH)[e], Jeffrey H. Samet (MD, MA, MPH)[b,f]

[a] *Alcohol and Drug Abuse Research Center (ADARC), McLean Hospital-Harvard Medical School, Boston, MA, 02478, USA*
[b] *Clinical Addiction Research and Education (CARE) Unit, Section of General Internal Medicine, Department of Medicine, Boston University School of Medicine and Boston Medical Center, Boston, MA, 02118, USA*
[c] *Department of Epidemiology and the Center to Prevent Alcohol Problems Among Young People, Boston University School of Public Health, Boston, MA, 02118, USA*
[d] *Department of Mathematics, Smith College, Northampton, MA, 01063, USA*
[e] *Data Coordinating Center, Boston University School of Public Health, Boston, MA, 02118, USA*
[f] *Department of Social and Behavioral Sciences, Boston University School of Public Health, Boston, MA, 02118, USA*

Received 8 July 2004; accepted 4 August 2004

Abstract

Introduction: Individuals with substance-related disorders are at increased risk for suicidal behavior. Identifying those at higher risk for suicide among this population is difficult and informed mainly on the basis of cross-sectional data.
Methods: We examined factors associated with drug-related suicidal behavior using multivariable regression analyses in a 2-year prospective study of 470 inpatients enrolled from an unlocked, detoxification unit. Suicidal behavior included suicidal ideation (SI) and suicide attempt (SA).
Results: Lifetime prevalence for SI was 28.5%, and for SA, 21.9%. During the 2-year follow-up, 19.9% of the sample endorsed suicidal ideation and 6.9% reported a suicide attempt. Correlates of lifetime suicidal behavior included younger age, female, Hispanic, greater depressive symptoms, past sexual abuse, and problem sedative or alcohol use. Factors associated with suicidal behavior at follow-up included past suicidal behavior, more depressive symptoms, and more frequent benzodiazepine and alcohol use. Cocaine and heroin use did not reach statistical significance.
Conclusions: Suicidal behavior is common among individuals with substance-related disorders. Differences in "suicide potential" may exist between drug categories with CNS depressants increasing the risk. These findings highlight the importance of addressing the recurrent 'suicide risk' of patients with substance-related disorders and regular monitoring for changes in depressive symptoms and drug use. Based on the prevalence and severity of this problem, the role of *universal* suicide screening of individuals with substance-related disorders merits greater attention.

Keywords: Suicide attempt; Suicidal ideation; Drug use; Depressive symptoms; Substance-related; Prospective cohort

1. Introduction

Developing effective suicide prevention remains a formidable challenge. Approximately 30,000 Americans are lost to suicide each year. Suicide is the third leading cause of death in individuals ages 15–24 (Centers for Disease Control and Prevention, 2004) and the 11th leading cause of death overall. Research has consistently demonstrated that

☆ Preliminary results were presented at the annual national meeting of the College on Problems of Drug Dependence (CPDD), 14–19 June 2003 in Bal Harbour, Florida and at NIDA/UR workshop on "Drug Abuse and Suicidal Behavior," in 27–28 March 2003, Bethesda, Maryland.

* Corresponding author. Present address: McLean Hospital, 115 Mill Street, Belmont, MA 02478, USA. Tel.: +1 617 855 2769; fax: +1 617 855 2519.
E-mail address: jwines@hms.harvard.edu (J.D. Wines Jr.).

0376-8716/$ – see front matter © 2004 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.drugalcdep.2004.08.004

a significant percentage of suicides and suicide attempts are alcohol or other drug-related (Mendelson and Rich, 1993; Moscicki, 1997). Although extensive scientific literature suggests that individuals with substance-related disorders are at increased risk for suicidal behavior, a vague knowledge of the clinical nature of the drug-suicidal behavior link hampers both risk-identification and prevention efforts.

A large body of evidence suggests that mood disorders and depressive symptoms increase the risk for suicidal behavior in subjects with substance-related disorders (Darke et al., 2004; Kosten and Rounsaville, 1988; Mendelson and Rich, 1993; Moscicki, 1997; Pages et al., 1997; Preuss et al., 2003; Roy, 2002; Roy, 2001). A retrospective study of patients presenting for an initial psychiatric evaluation, found that alcoholics with major depression exhibited 59% more suicidality than patients with major depression only (Cornelius et al., 1995). Recent cocaine use by depressed alcoholics may increase suicidality further (Cornelius et al., 1998). Another retrospective study of 891 psychiatric inpatients with a primary diagnosis of non-psychotic, unipolar depression, concluded that high levels of current alcohol/drug use were associated with increased suicidal ideation (Pages et al., 1997).

Despite advances, there are still substantial gaps in our current understanding of drug-related suicidal behavior. Although the risk of suicidal behavior associated with alcohol intoxication has been supported by numerous studies (Borges and Rosovsky, 1996; Mayfield and Montgomery, 1972; Suokas and Lonnqvist, 1995) relatively little is known about the impact of different drug categories, drug combinations, substance-induced effects and self-medication on suicidal behavior (Aharonovich et al., 2002; Preuss et al., 2003). Recently, retrospective studies have begun to describe characteristics of suicidal behavior by individuals dependent on different types of drugs (e.g., cocaine dependence, opioid dependence) (Darke et al., 2004; Garlow et al., 2003; Roy, 2002; Roy, 2001). Thus, it is not surprising that prospective studies examining the relationship between changes in drug use, depressive symptoms and suicidal behavior are scant.

The purpose of this prospective cohort study was to more clearly define the phenomenology of drug-related suicidal behavior by identifying factors associated with non-fatal suicidal behavior (suicidal ideation and suicide attempt) in a sample of patients initially admitted for detoxification. Three hypotheses were tested: (1) differences in suicidal behavior will exist across drug categories; (2) past suicide attempt will be associated with suicide attempt during follow-up; and (3) more depressive symptoms and/or more frequent drug use at follow-up will be significantly and independently associated with suicidal behavior.

2. Methods

2.1. Design

The health evaluation and linkage to primary care (HELP) study was a randomized clinical trial (RCT) testing the effectiveness of a multidisciplinary health intervention to link alcohol and other drug-dependent individuals to primary care. A detailed description of the HELP study RCT has been previously reported (Samet et al., 2003). After providing written informed consent, eligible subjects were enrolled in this RCT. In this study, we consider the prospective follow-up of this cohort focusing on suicidal behavior. We attempted to follow subjects every 6 months for 2 years. The Institutional Review Board at Boston University Medical Center approved the study. Additional privacy protection was secured by the issuance of a Certificate of Confidentiality by the Department of Health and Human Services.

2.2. Subjects

Adults admitted to an unlocked, inpatient detoxification unit in Boston, Massachusetts, between 1 February 1997 and 1 April 1999 were screened for the HELP study. Most clients admitted to this inpatient facility were either uninsured or would qualify for Medicaid benefits. Subject eligibility criteria were the following: (1) admitted for inpatient detoxification; (2) identified alcohol, cocaine, or heroin as first or second drug of choice; (3) did not have a primary care physician; (4) over 17 years old; (5) not pregnant; (6) fluent in English or Spanish; (7) lived in the Boston Medical Center catchment area (Boston metropolitan area, primarily inner city) or were homeless; (8) not planning to leave the Boston area within the next 12 months; (9) able to provide three contact names to assist with follow-up tracking; and (10) scored >20 on the mini-mental state exam (Folstein et al., 1975). Supermarket certificates, US$ 20 for the initial evaluation and US$ 30 per follow-up interview, were used to compensate participants for their time.

2.3. Assessments and measures

Trained research associates administered standard interviews to subjects at study entry (i.e., the index admission for detoxification) and during follow-up interviews at the Boston University General Clinical Research Center or via telephone. During detoxification, subjects were interviewed only after the resolution of acute withdrawal symptoms. For the Spanish interview standardized scales in Spanish were used when available; the remainder of the questionnaire was translated from English into Spanish, back-translated to check for accuracy, and then corrected. Assessments pertinent to the current investigation included: demographics (age, gender, and race/ethnicity); homeless status; history of any sexual and/or physical abuse; Addiction Severity Index (ASI) alcohol and drug sub-scales (McLellan et al., 1992); depressive symptoms as measured by the Center for Epidemiologic Studies Depression (CES-D) scale (Radloff, 1977); physical health status as measured by the Short Form Health Survey, physical component summary (SF-36-PCS) (Ware, 1993); and indicators of social support from friends and family assessed by the Perceived Social Support Scale (PSS)

(Procidano and Heller, 1983). Homelessness, a dichotomous variable, was defined as having lived in a homeless shelter or on the street for 1 or more days during the previous 6 months.

2.3.1. Problem Drug categories at detoxification admission

Problem Drug categories were constructed from the 11 drugs (alcohol, heroin, methadone, other opioids/analgesics, barbiturates, sedatives/hypnotics/tranquilizers, cocaine, amphetamine, cannabis, hallucinogens, and inhalants) listed in the ASI at the index admission for detoxification. In the absence of formal drug diagnoses (abuse or dependence), frequency of drug use was used as a marker for excessive or heavy use. *Problem Drug* for a particular substance was defined as use $\geq$3 times per week (for alcohol "use" means to intoxication or $\geq$3 drinks) for a year or more, or 5 or more days of use in the past 30 days (definitions based on the ASI and outcome definitions used in treatment studies) (McLellan et al., 1992; O'Malley et al., 1992; Volpicelli et al., 1992). Theoretical considerations were used to combine drugs with similar pharmacologic properties for analysis, creating four *Problem Drug* categories: alcohol; opioids (heroin, methadone and other opioids/analgesics); sedatives (barbiturates and sedatives/hypnotics/tranquilizers); and stimulants (cocaine and amphetamines). Indicator variables for each *Problem Drug* category were produced by combining the relevant drugs listed in the ASI. For example, if a subject met the *Problem Drug* definition for any of the three opioid drugs (heroin, methadone and other opioids/analgesics), then the subject would be considered to have problem opioid use (Opioid *Problem Drug*), an indicator of frequent use and a proxy for opioid abuse/dependence. Also, a continuous variable representing the total number of the eleven drugs listed on the ASI that met the *Problem Drug* criteria was created.

2.3.2. Drug Use after detoxification (follow-up)

Drug Use was defined as the number of days used in the past 30 days as recorded by the ASI (range, 0–30 days) during follow-up. For *Problem Drug* categories consisting of more than one drug (i.e., opioids, sedatives, stimulants), the single most commonly used drug by the cohort at study entry (i.e., heroin, benzodiazepines, cocaine) was utilized. This 'one drug per category' approach was taken in part to minimize ambiguity regarding the number of days used, as it cannot be determined from the ASI whether drugs used within a category are used on the same day or on different days. It is not surprising that heroin, cocaine and benzodiazepines were the most common drug used in their respective categories since the HELP study inclusion criteria required that heroin, cocaine or alcohol be a drug of choice for the participant and benzodiazepine use is much more common than barbiturate use (Substance Abuse and Mental Health Services Administration, 2003). Specifically, Alcohol *Use* was defined as the number of days of any alcohol use; Heroin, Benzodiazepine, and Cocaine *Use* were defined as, the number of days of each drug used, respectively. In addition, a continuous variable representing the total number of the eleven drugs listed on the ASI with any use in the past 30 days was created as a marker for polydrug use.

2.4. Primary outcomes

The primary outcomes for this study were suicidal ideation and suicidal attempt. Suicidal behavior (ideation and attempt) was assessed by two questions from the Addiction Severity Index (ASI): "Have you attempted suicide? (include actual suicidal gestures or attempts)" and "Have you experienced serious thoughts of suicide? (patient seriously considered a plan for taking his/her life)." Suicidal behavior was assessed at study entry, 6, 12, 18 and 24 months. Lifetime suicidal behavior was ascertained at study enrollment by asking subjects if they ever had suicidal ideation or suicide attempt. Suicidal behavior after detoxification was assessed by asking subjects if they had suicidal ideation or suicide attempt during their follow-up interval.

2.5. Statistical analyses

Logistic regression models were used to predict lifetime (past) suicidal behavior, using *Problem Drug* categories as covariates, while controlling for possible confounders. Proportional hazards regression models were used to predict time to first report of suicidal behavior during the follow-up period, controlling for recent drug use, recent depressive symptoms and other possible confounders. Updated information about drug use and depressive symptoms at follow-up timepoints (every 6 months) was used in the time-to-event models. Kaplan–Meier estimates were used to determine the proportion of subjects exhibiting suicidal behavior during the 2-year follow-up interval. Because different risk factors have been reported for suicidal ideation and suicide attempt, separate statistical analyses were performed for each of these outcomes.

Variable selection was based on previous scientific studies and clinical experience. To test the first hypothesis we used data collected at study entry only. Models for the cross-sectional logistic regression analysis included the variables age, gender, race, homeless status, SF-36-physical component summary (PCS), perceived social support (Family), perceived social support (Friends), history of physical or sexual abuse, depressive symptoms (CES-D), and *Problem Drug* category. Secondary analyses were conducted entering only one *Problem Drug* category per model rather than all four simultaneously. Similarly, a single continuous variable representing the total number of drugs listed on the ASI fulfilling the *Problem Drug* criteria was also tested in the models.

Data from follow-up were used to test the second and third hypotheses. In the interest of parsimony and because of a small number of outcomes, only variables that were significantly associated with suicidal behavior from the cross-sectional logistic regression models or considered clinically essential were entered into the longitudinal models. Models

for the longitudinal analysis contained the following variables from the initial assessment: age, gender, Hispanic ethnicity, past sexual abuse, and past suicidal ideation or attempt. These models also contained recent (past 7 days) depressive symptoms and recent (past 30 days) *Drug Use*. These latter two variables were assessed during follow-up and entered as time-varying covariates. These models were also fit without depressive symptoms as a covariate because of concerns about collinearity with drug use. Secondary analyses were conducted entering a single continuous variable representing the total number of drugs listed on the ASI with any use in the past 30 days into the models. A $p < 0.05$ was considered statistically significant and SAS/STAT software version 8.2 (Cary, NC) was used to perform all analyses.

3. Results

3.1. Sample characteristics

The majority (76%) of the 470 subjects enrolled in the study were male. The mean (SD) age was 36 (8). Almost half (46%) were Black, 37% White, and 11% were Hispanic. During the 6 months prior to the index detoxification, nearly half (47%) of the population was homeless at least one night and 39% were unemployed. A lifetime history of sexual abuse with or without physical abuse was reported by 29% and 43% had a history of physical abuse only. This cohort was characterized by the following scale scores {mean (SD, range)}: SF-36 physical component summary (PCS) 48 (11, 14–75); CES-D 33 (12, 1–60), with 90% ≥16, a value associated with clinically significant symptoms (Boyd et al., 1982); and perceived social support for family and friends 7 (5, 0–14) and 7 (4, 0–14), respectively. The frequency of *Problem Drug* categories was as follows: (1) alcohol: 86%, (2) opioid: 43%, (3) sedative: 23%, and (4) stimulant: 77%. Two subjects died prior to follow-up, and 85% (400/468) of the remaining subjects were assessed at least once during the 2-year follow-up interval. Participants completed a mean of 17.5 months of follow-up (median: 24 months). Further details about the cohort and the follow-up experience have been previously published (Samet et al., 2003).

3.2. Lifetime suicidal behavior

The prevalence of lifetime suicidal ideation and suicide attempt was 28.5% and 21.9%, respectively (Table 1, upper panel). The following characteristics were associated with suicide ideators, as compared to non-ideators: younger age; Hispanic ethnicity; less social support from family; a history of sexual abuse; more depressive symptoms, and Alcohol *Problem Drug* (Table 2). The odds of having a history of suicidal ideation for patients with Alcohol *Problem Drug* category were 2.66 times the odds of those without Alcohol *Problem Drug* designation (OR: 2.66, 95% CI: 1.15–6.14). Subjects reporting Sedative *Problem Drug* were more likely to report ideation at borderline statistical significance ($p = 0.09$). Comparing suicide attempters to non-attempters, attempters had the following characteristics: younger age; female; Hispanic ethnicity; a history of sexual abuse; more depressive symptoms; and Sedative *Problem Drug* (Table 2). More specifically, patients with Sedative *Problem Drug*, as compared to those without Sedative *Problem Drug*, had 2.65 the odds of a history of suicide attempt (OR: 2.65, 95% CI: 1.35–5.21). At borderline statistical significance, subjects reporting Alcohol *Problem Drug* were more likely to report an attempt ($p = 0.10$). Data were re-analyzed using only one *Problem Drug* category per model and these results were similar to models containing all four *Problem Drug* categories. The total number of drugs listed on the ASI that fulfilled the *Problem Drug* criteria was associated with suicidal ideation (OR: 1.25, 95% CI: 1.11–1.42) and suicide attempt (OR: 1.22, 95% CI: 1.07–138). Here the OR represents the change for each additional *Problem Drug*.

3.3. Suicidal behavior after detoxification

Kaplan–Meier estimates demonstrated that approximately 20% of the sample reported suicidal ideation at follow-up; 6.9% reported suicide attempt (Table 1, lower panel). During the 2-year follow-up, 69 individuals endorsed suicidal ideation, and 23 individuals acknowledged a suicide attempt. Almost half (47%) of subjects with lifetime suicidal ideation reported at study entry exhibited suicidal ideation after detoxification. Likewise, nearly one-fourth (24%) of those with a

Table 1
Suicidal ideation (SI) or attempt (SA): lifetime and during 24 months after detoxification

Lifetime prevalence of suicidal behavior (assessed during detoxification) ($n = 470$)			
Suicidal ideation	28.5% (100/470)		
Suicide attempt	21.9% (85/470)		
Suicidal behavior during 24-months after detoxification[a]			
Suicidal ideation	19.9% (15.7–24.1) (total sample SI, $n = 396$[b])	46.5% (36.7–56.3) (with prior SI, $n = 121$[c])	8.4% (4.8–12.0) (w/o prior SI, $n = 275$[c])
Suicide attempt	6.9% (4.2–9.6) (total sample SA, $n = 396$[b])	24.1% (14.1–34.1) (with prior SA, $n = 91$[c])	2.3% (0.5–4.1) (w/o prior SA, $n = 305$[c])

95% confidence interval in parenthesis (%); SI: suicidal ideation; SA: suicide attempt.
[a] Kaplan–Meier estimates.
[b] With at least one follow-up.
[c] With or without suicidal behavior prior to detoxification and at least one follow-up.

Table 2
Factors associated with lifetime suicidal behavior (multivariable logistic regression)[a] (n = 470)

	Suicidal ideation			Suicidal attempt		
Characteristic:	OR	95% CI OR	p-value	OR	95% CI OR	p-value
Age	0.97	0.94, 1.00	0.04	0.96	0.93, 1.00	0.04
Female	1.61	0.87, 2.98	0.13	2.15	1.13, 4.06	0.02
Race/ethnicity						
White[b]	1.39	0.78, 2.47	0.27	1.40	0.73, 2.67	0.31
Hispanic[b]	3.00	1.37, 6.61	0.006	3.77	1.64, 8.71	0.002
Other[b]	1.12	0.40, 3.12	0.83	1.40	0.48, 4.12	0.54
Homeless	1.36	0.83, 2.23	0.23	0.88	0.51, 1.51	0.63
PCS	1.00	0.97, 1.02	0.84	1.00	0.97, 1.02	0.90
Social support (family)	0.93	0.88, 0.98	0.01	0.96	0.91, 1.02	0.21
Social support (friends)	1.02	0.95, 1.08	0.65	0.99	0.92, 1.06	0.67
Depressive symptoms	1.06	1.04, 1.09	<0.001	1.05	1.03, 1.08	<0.001
Past sexual abuse	2.60	1.32, 5.12	0.006	3.36	1.56, 7.23	0.002
Physical abuse only	1.40	0.75, 2.63	0.29	1.70	0.81, 3.55	0.16
Problem Drug[c]						
Alcohol	2.66	1.15, 6.14	0.02	2.11	0.87, 5.11	0.10
Opioids	1.35	0.77, 2.38	0.29	0.63	0.33, 1.18	0.15
Sedatives	1.70	0.92, 3.16	0.09	2.65	1.35, 5.21	0.005
Stimulants	1.33	0.73, 2.45	0.35	1.49	0.76, 2.89	0.25

OR: odds ratio; CI: confidence interval.
[a] Models control for all covariates listed above.
[b] Black as reference group.
[c] See Section 2 for definition of *Problem Drug*.

prior attempt, attempted suicide during the ensuing 2 years. For subjects with no prior suicidal behavior, the 24-month incidence of suicidal ideation was 8.4% and the incidence of suicide attempt was 2.3% (Table 1, lower panel).

In a multivariable analysis, prior suicidal ideation was significantly associated with future ideation (HR: 4.75, 95% CI: 2.75–8.21) (see Table 3). Younger age was associated with suicidal ideation at borderline statistical significance (p = 0.09). Recent depressive symptoms were significantly associated with suicidal ideation (HR 1.06, 95% CI: 1.03–1.08 for each point increase in the CES-D score). *Drug Use* was not significantly associated with suicidal ideation when depressive symptoms were included in the model. However, in a model that did not adjust for depressive symptoms, more frequent use of benzodiazepines (HR: 1.05, 95% CI: 1.01–1.08) and alcohol (HR: 1.03, 95% CI: 1.01–1.06) were associated with suicidal ideation. The HR for benzodiazepine and alcohol use represents the HR for each additional day of use during the past month. Thus 7 days of benzodiazepine use in the past month increased the hazard of suicidal ideation by 36% (HR: 1.36) and each week of alcohol use in the past month increased the hazard of suicidal ideation by 25% (HR: 1.25).

In a separate multivariable analysis, previous history of suicide attempt was significantly associated with future at-

Table 3
Factors associated with suicidal ideation after detoxification (multivariable proportional hazards models with and without depressive symptoms)[a]

	With depressive symptoms			Without depressive symptoms		
Characteristic:	HR	95% CI HR	p-value	HR	95% CI HR	p-value
Prior suicidal ideation	4.75	2.75, 8.21	<0.001	6.09	3.51, 10.54	<0.001
Age	0.97	0.94, 1.01	0.09	0.99	0.95, 1.02	0.39
Female	1.19	0.67, 2.10	0.56	1.54	0.88, 2.68	0.13
Hispanic vs. non-Hispanic	1.42	0.70, 2.87	0.33	1.49	0.75, 2.95	0.26
Depressive symptoms[b]	1.06	1.03, 1.08	<0.001	–	–	–
Any sexual abuse	1.64	0.95, 2.83	0.08	1.47	0.85, 2.54	0.17
Drug Use[b]						
Alcohol	1.02	0.99, 1.04	0.21	1.03	1.01, 1.06	0.01
Heroin	1.00	0.98, 1.03	0.78	1.02	0.99, 1.05	0.18
Benzodiazepines	1.02	0.98, 1.06	0.38	1.05	1.01, 1.08	0.01
Cocaine	0.99	0.95, 1.02	0.44	1.00	0.97, 1.04	0.88

HR: hazard ratio; CI: confidence interval.
[a] Models control for all covariates above
[b] Time varying covariates (depressive symptoms: past 7 days and *Drug Use*: # days past 30)

Table 4
Factors associated with suicide attempt after detoxification (multivariable proportional hazards models with and without depressive symptoms)[a]

	With depressive symptoms			Without depressive symptoms		
Characteristic:	HR	95% CI HR	*p*-value	HR	95% CI HR	*p*-value
History of suicidal attempt	6.12	2.23, 16.79	<0.001	8.81	3.20, 24.20	<0.001
Age	0.97	0.91, 1.03	0.31	0.99	0.94, 1.05	0.75
Female	0.97	0.36, 2.61	0.95	1.22	0.46, 3.22	0.69
Hispanic vs. non-Hispanic	2.67	0.84, 8.66	0.1	2.60	0.86, 7.88	0.09
Depressive symptoms[b]	1.06	1.03, 1.10	<0.001	–	–	–
Any sexual abuse	1.56	0.58, 4.16	0.38	1.38	0.52, 3.69	0.52
Drug Use[b]						
Alcohol	1.00	0.95, 1.05	0.99	1.03	0.99, 1.08	0.18
Heroin	0.94	0.85, 1.03	0.16	0.96	0.88, 1.05	0.37
Benzodiazepines	1.00	0.95, 1.06	0.98	1.02	0.97, 1.08	0.42
Cocaine	0.96	0.89, 1.05	0.37	0.97	0.90, 1.05	0.49

HR: hazard ratio; CI: confidence interval.
[a] Models control for all covariates above.
[b] Time varying covariates (depressive symptoms: past 7 days and *Drug Use*: # days past 30).

tempt (HR: 6.12, 95% CI: 2.23–16.79) (Table 4). Recent depressive symptoms were significantly associated with attempts (HR: 1.06, 95% CI: 1.03–1.10). There were no significant associations between frequency of recent *Drug Use* (alcohol, heroin, benzodiazepines or cocaine) and suicide attempts, whether or not depressive symptoms were in the model (Table 4).

In secondary analyses, we replaced the four *Drug Use* variables with the four *Problem Drug* categories. None of the four *Problem Drug* categories were significantly associated with either suicidal ideation or attempt. We also ran the models with a single continuous variable representing the total number of drugs with any use in the past 30 days. For suicidal ideation, the total number of drugs was not significantly associated in the model with CES-D (HR: 1.09, 95% CI: 0.92–1.28). However, it was a significant predictor of time to ideation (HR: 1.28, 95% CI: 1.11–1.48) when CES-D was not included. For suicide attempts, the total number of drugs was not significantly associated when CES-D was omitted (HR: 0.92, 95% CI: 0.69–1.22) but was borderline significantly associated ($p = 0.06$) in a non-hypothesized direction when CES-D was included (HR: 0.73, 95% CI: 0.52–1.01).

4. Discussion

Suicidal behavior is a common dimension of life for patients requiring detoxification. Such life threatening thoughts and actions merit study so as to enable recognition of an individual's suicidal risk and facilitate appropriate action to minimize consequences. The major findings of this study are fourfold: (1) differences for lifetime suicidal behavior exist between drug categories; (2) prior suicidal behavior is an important risk factor for future suicidal behavior; (3) more depressive symptoms at follow-up are associated with suicidal behavior; and (4) more frequent use of benzodiazepines and alcohol is associated with follow-up suicidal ideation when depressive symptoms are not controlled for in a multivariable model. This study uniquely adds to the literature an examination of suicidal behavior in a prospective study of a large, ethnically diverse cohort discharged from a detoxification unit and demonstrates that more depressive symptoms and more frequent drug use (benzodiazepines and alcohol) at follow-up are associated with suicidal ideation.

Like earlier reports, suicidal behavior is common among patients admitted with drug-related disorders (Pages et al., 1997; Roy, 2002; Roy, 2001). One-fifth of patients acknowledged ideation at follow-up; 6.9% reported a suicide attempt over a 2-year follow-up. The incidence of ideation and attempt in those with no past suicidal behavior was 8.4% and 2.3%, respectively—rates lower than the 20.6% for "new suicidal intent" and 5.5% for new onset suicide attempt reported from a 2.5 year follow-up of opioid-dependent patients (Kosten and Rounsaville, 1988). These discrepancies may be explained in part by differences in the sample populations (opioids versus a variety of substances used) and operational definitions of suicidal behavior.

In our sample, a history of suicidal ideation and attempt were strong predictors of future ideation and attempt, respectively. A recent 5-year follow-up of 1237 alcoholics also demonstrated that previous suicide attempt was strongly associated with future attempts (Preuss et al., 2003). The replication of this result in a prospective sample of subjects using a variety of drugs has significant clinical importance given that a history of attempted suicide may increase the likelihood of suicide completion over 30-fold (Harris and Barraclough, 1997).

Our data suggest CNS depressants like benzodiazepines, barbiturates and alcohol elevate the risk for suicidal behavior. Several potential mechanisms for alcohol triggering suicidal behavior have been suggested including alcohol's ability to disinhibit behavior, increase depression, facilitate aggression, and narrow attention (Brown and Schuckit, 1988; Hufford, 2001). Other CNS depressants may work through similar mechanisms, although empirical data are lacking. The small number of studies that have investigated whether differ-

ences in the "suicide potential" exist between drug categories yielded contradictory findings. In a study of 155 "polydrug abusers," Ward and Schuckit (1980) reported that subjects using primarily sedative/hypnotic drugs had more serious suicide attempts (Ward and Schuckit, 1980). In addition, a cross-sectional study of 2051 patients with drug-related disorders in Norway found that daily users of tranquilizers and alcohol were significantly more likely than daily users of stimulant or opioid drugs to attempt suicide (Rossow and Lauritzen, 1999). In contrast, Borges et al., 2000 analyzing data from the National Comorbidity Survey, reported similar risk for suicide attempt across most drug types, concluding that the number of different drugs used was more important than the type of drug used (Borges et al., 2000). The number of drugs in our study was significantly associated with lifetime suicidal behavior and suicidal ideation after detoxification. The differences in these results from the National Comorbidity Survey may be due to methodology (i.e., community versus treatment study populations).

Data obtained from our subjects after discharge from the detoxification unit reveal an interesting, yet complex, relationship between depressive symptoms, drug use and suicidal behavior. In addition to drug use, depressive symptoms can result from a variety of sources including mood disorder, adjustment disorder, bereavement, and medical illness (DSM-IV, 1994; Raimo and Schuckit, 1998). This could partially explain the robust association of depressive symptoms at follow-up with both suicidal ideation and suicide attempt, regardless of controlling for drug use in the multivariable model. For drug use, no association was found with follow-up suicide attempt, whether or not depressive symptoms were in the multivariable model. However, more frequent drug use (alcohol and benzodiazepines) was associated with follow-up suicidal ideation when the model did not include depressive symptoms, suggesting collinearity between drug use and depressive symptoms. It is plausible that drug use and depressive symptoms mediate each other—e.g., increasing substance use causes increased depressive symptoms (substance-induced depression) (Brown and Schuckit, 1988; Ries et al., 2001). Alternatively, increasing depressive symptoms may lead directly to increasing substance use (self-medication) (Khantzian, 1985) (Fig. 1). Drug use and depressive symptoms may lie on the same mechanistic pathway for suicidal ideation, a hypothesis that is consistent with the data, however, the directionality of this relationship cannot be determined with the current study design.

Our findings are consistent with a recent prospective study analyzing data from Project MATCH which found that both drinking and depressive symptoms were associated with suicidal ideation in alcoholics (Conner et al., 2003), although direct comparison with our study is limited partly by the smaller sample size which precludes stratification by gender and analysis of different drinking patterns. Future research should focus on the precise sequence of events for drug use and depressive symptoms as well as studying the association between suicidal behaviors and more sensitive measures of

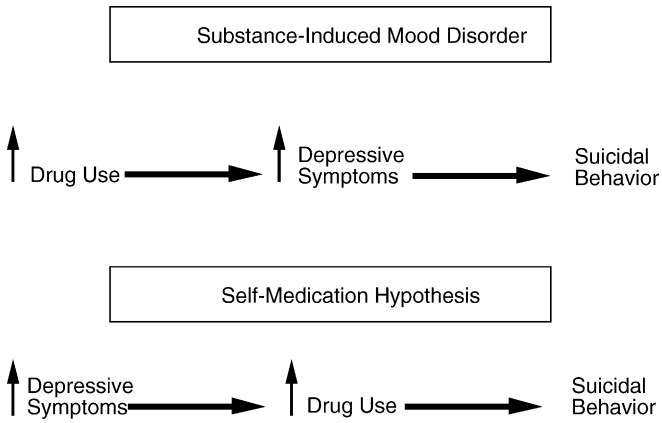


Fig. 1. Models of the relationship between drug use, depressive symptoms, and suicidal behavior.

frequency (e.g., number of times per day used) and quantity (e.g., grams of cocaine, bags of heroin, amount of money spent) of *Drug Use* that have the potential to reveal distinct use patterns as well as more accurate estimates of dose.

Age, female gender, and history of sexual abuse are well-established risk factors for suicidal behavior (Borges et al., 2000; Moscicki, 1997; Roy, 2001, 2002). In our data, younger age, female gender and history of sexual abuse were associated with lifetime suicidal behavior, but not suicidal behavior at follow-up possibly due to insufficient power. The increased risk for suicidal behavior by Hispanics has been reported by some investigators but not by others (Canino and Roberts, 2001; Garofalo et al., 1999; Oquendo et al., 2001; Tortolero and Roberts, 2001). Clearly, this potential ethnic disparity in suicidal behavior merits further investigation.

This study has some important limitations. Although widely accepted and validated instruments were used, they assess different time periods (i.e., depressive symptoms (CES-D): one week; drug use (ASI): 1 month). Therefore, we cannot draw firm conclusions regarding the temporal relationship between depressive symptoms, drug use and suicidal behavior. As the study was not powered to detect multivariable associations for suicidal behavior, the relatively low base rate for suicidal behavior after study entry, especially suicide attempts, may have resulted in Type II errors. Diagnostic instruments (e.g., SCID) were not used to obtain DSM-IV defined substance abuse, substance dependence, or depression making comparisons with diagnosed populations more difficult. However, the stringent admission criteria for limited public residential detoxification beds as well as the severity of results from other addiction measures in the study (e.g., ASI scores) strongly suggest that all subjects met criteria for substance dependence. Because our study recruited subjects from an unlocked detoxification unit, acutely suicidal patients in need of detoxification were most likely admitted to a more secure facility (i.e., a locked unit). Thus our results probably underestimate the extent of suicidality in this population. Finally, results from this investigation may not generalize to dissimilar populations (e.g., patients whose drug of choice is

not alcohol, heroin or cocaine; individuals not in treatment; adolescents).

There also are several strengths of the study: the prospective study design; the diverse cohort, both ethnically and in terms of drug preferences; administration of standardized instruments by trained research associates; and application of survival methods that account for censoring and time-varying covariates.

In summary, this study highlights the importance of suicidal behavior in a "captive" population, residential detoxification patients, one that may benefit from its recognition and treatment. Recently, Darke et al. (2004) recommended that individuals with heroin dependence receive routine screening for suicidal behavior and depression. The presence of such substantial disease prevalence and incidence as well as the potential for targeted intervention, gives urgency to the consideration of recurrent *universal* 'suicide-risk' screening of all patients with substance-related disorders. The feasibility and benefit of such a policy for regular monitoring of suicidal behavior requires attention. The focus on a detoxification population as a contact with the health care system that could be utilized to address suicide risk is pragmatic. Our findings also suggest potential clinical utility of routine interval testing with standardized instruments for depressive symptoms (e.g., CES-D), drug use (e.g., ASI) and suicidal behavior. Despite similarities between risk factors for suicidal ideation and suicide attempt, it is important to emphasize that their relation to completed suicide is uncertain. Additional studies are necessary to further characterize the natural history of drug-related suicidal behavior and shed light on the complex interplay of the drug-individual-environment triad. Understanding these relationships is crucial for developing drug-sensitive, suicide risk-assessment instruments, innovative intervention strategies and enhanced suicide prevention.

Acknowledgments

This work was supported by grants from the National Institute on Drug Abuse. "Drug-Related Suicidal and/or Homicidal Behavior," K23-DA4550045 (JW), "Enhanced Linkage of Drug Abusers to Primary Medical Care," R01-DA10019 and the National Institute on Alcohol Abuse and Alcoholism, "Enhanced Linkage of Alcohol Abusers to Primary Care," R01-AA10870. This work was also supported in part by a General Clinical Research Center grant MO1-RR00533 from the National Center for Research Resources.

References

Aharonovich, E., Liu, X., Nunes, E., Hasin, D.S., 2002. Suicide attempts in substance abusers: effects of major depression in relation to substance use disorders. Am. J. Psychiatry 159, 1600–1602.

Borges, G., Rosovsky, H., 1996. Suicide attempts and alcohol consumption in an emergency room sample. J. Stud. Alcohol 57, 543–548.

Borges, G., Walters, E.E., Kessler, R.C., 2000. Associations of substance use, abuse, and dependence with subsequent suicidal behavior. Am. J. Epidemiol. 151, 781–789.

Brown, S.A., Schuckit, M.A., 1988. Changes in depression among abstinent alcoholics. J. Stud. Alcohol 49, 412–417.

Boyd, J.H., Weissman, M.M., Thompson, W.D., Myers, J.K., 1982. Screening for depression in a community sample. Understanding the discrepancies between depression symptom and diagnostic sales. Arch. Gen. Psychiatry 39, 1195–1200.

Canino, G., Roberts, R.E., 2001. Suicidal behavior among Latino youth. Suicide Life Threat. Behav. 31 (Suppl.), 122–131.

Centers for Disease Control and Prevention, 2004. Injury Fact Sheets: Suicide in the United States. National Center for Injury Prevention and Control, Centers for Disease Control and Prevention (producer). http://www.cdc.gov/ncipc/factsheets/suifacts.htm (accessed 4/30/04).

Conner, K.R., Li, Y., Meldrum, S., Duberstein, P.R., Conwell, Y., 2003. The role of drinking in suicidal ideation: analyses of Project MATCH data. J. Stud. Alcohol 64, 402–408.

Cornelius, J.R., Salloum, I.M., Mezzich, J., Cornelius, M.D., Fabrega Jr., H., Ehler, J.G., Ulrich, R.F., Thase, M.E., Mann, J.J., 1995. Disproportionate suicidality in patients with comorbid major depression and alcoholism. Am. J. Psychiatry 152, 358–364.

Cornelius, J.R., Thase, M.E., Salloum, I.M., Cornelius, M.D., Black, A., Mann, J.J., 1998. Cocaine use associated with increased suicidal behavior in depressed alcoholics. Addict. Behav. 23, 119–121.

Darke, S., Ross, J., Lynskey, M., Teesson, M., 2004. Attempted suicide among entrants to three treatment modalities for heroin dependence in the Australian Treatment Outcome Study (ATOS): prevalence and risk factors. Drug Alcohol. Depend. 73, 1–10.

DSM-IV, 1994. Diagnostic and Statistical Manual of Mental Disorders. American Psychiatric Association, Washington, DC.

Folstein, M.F., Folstein, S.E., McHugh, P.R., 1975. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J. Psychiatr. Res. 12, 189–198.

Garlow, S.J., Purselle, D., D'Orio, B., 2003. Cocaine use disorders and suicidal ideation. Drug. Alcohol. Depend. 70, 101–104.

Garofalo, R., Wolf, R.C., Wissow, L.S., Woods, E.R., Goodman, E., 1999. Sexual orientation and risk of suicide attempts among a representative sample of youth. Arch. Pediatr. Adolesc. Med. 153, 487–493.

Harris, E.C., Barraclough, B., 1997. Suicide as an outcome for mental disorders. A meta-analysis. Br. J. Psychiatry 170, 205–228.

Hufford, M.R., 2001. Alcohol and suicidal behavior. Clin. Psychol. Rev. 21, 797–811.

Khantzian, E.J., 1985. The self-medication hypothesis of addictive disorders: focus on heroin and cocaine dependence. Am. J. Psychiatry 142, 1259–1264.

Kosten, T.R., Rounsaville, B.J., 1988. Suicidality among opioid addicts: 2.5 year follow-up. Am. J. Drug. Alcohol. Abuse. 14, 357–369.

Mayfield, D.G., Montgomery, D., 1972. Alcoholism, alcohol intoxication, and suicide attempts. Arch. Gen. Psychiatry 27, 349–353.

McLellan, A.T., Kushner, H., Metzger, D., Peters, R., Smith, I., Grissom, G., Pettinati, H., Argeriou, M., 1992. The fifth edition of the addiction severity index. J. Subst. Abuse Treat. 9, 199–213.

Mendelson, W.B., Rich, C.L., 1993. Sedatives and suicide: the San Diego study. Acta Psychiatr. Scand. 88, 337–341.

Moscicki, E.K., 1997. Identification of suicide risk factors using epidemiologic studies. Psychiatr. Clin. North Am. 20, 499–517.

O'Malley, S.S., Jaffe, A.J., Chang, G., Schottenfeld, R.S., Meyer, R.E., Rounsaville, B., 1992. Naltrexone and coping skills therapy for alcohol dependence. A controlled study. Arch. Gen. Psychiatry 49, 881–887.

Oquendo, M.A., Ellis, S.P., Greenwald, S., Malone, K.M., Weissman, M.M., Mann, J.J., 2001. Ethnic and sex differences in suicide rates relative to major depression in the United States. Am. J. Psychiatry 158, 1652–1658.

Pages, K.P., Russo, J.E., Roy-Byrne, P.P., Ries, R.K., Cowley, D.S., 1997. Determinants of suicidal ideation: the role of substance use disorders. J. Clin. Psychiatry 58, 510–515, quiz 516–517.

Preuss, U.W., Schuckit, M.A., Smith, T.L., Danko, G.P., Bucholz, K.K., Hesselbrock, M.N., Hesselbrock, V., Kramer, J.R., 2003. Predictors and correlates of suicide attempts over 5 years in 1,237 alcohol-dependent men and women. Am. J. Psychiatry 160, 56–63.

Procidano, M.E., Heller, K., 1983. Measures of perceived social support from friends and from family: three validation studies. Am. J. Community Psychol. 11, 1–24.

Radloff, L.S., 1977. The CES-D scale: a self-report depression scale for research in the general population. App. Psych. Measurement 1, 385–401.

Raimo, E.B., Schuckit, M.A., 1998. Alcohol dependence and mood disorders. Addict. Behav. 23, 933–946.

Ries, R.K., Demirsoy, A., Russo, J.E., Barrett, J., Roy-Byrne, P.P., 2001. Reliability and clinical utility of DSM-IV substance-induced psychiatric disorders in acute psychiatric inpatients. Am. J. Addict. 10, 308–318.

Roy, A., 2001. Characteristics of cocaine-dependent patients who attempt suicide. Am. J. Psychiatry 158, 1215–1219.

Roy, A., 2002. Characteristics of opiate dependent patients who attempt suicide. J. Clin. Psychiatry 63, 403–407.

Samet, J.H., Larson, M.J., Horton, N.J., Doyle, K., Winter, M., Saitz, R., 2003. Linking alcohol- and drug-dependent adults to primary medical care: a randomized controlled trial of a multidisciplinary health intervention in a detoxification unit. Addiction 98, 509–516.

Substance Abuse and Mental Health Services Administration (SAMHSA), 2003. Office of Applied Studies, Nonmedical use of prescription-type drugs among youths and young adults, The NHSDA Report. US Department of Health and Human Services, Rockland, MD.

Suokas, J., Lonnqvist, J., 1995. Suicide attempts in which alcohol is involved: a special group in general hospital emergency rooms. Acta Psychiatr. Scand. 91, 36–40.

Tortolero, S.R., Roberts, R.E., 2001. Differences in nonfatal suicide behaviors among Mexican and European American middle school children. Suicide Life Threat. Behav. 31, 214–223.

Volpicelli, J.R., Alterman, A.I., Hayashida, M., O'Brien, C.P., 1992. Naltrexone in the treatment of alcohol dependence. Arch. Gen. Psychiatry 49, 876–880.

Ward, N.G., Schuckit, M.A., 1980. Factors associated with suicidal behavior in polydrug abusers. J. Clin. Psychiatry 41, 379–385.

Ware, J.E., 1993. SF-36 Health Survey. Manual and Interpretation Guide. The Health Institute, Boston, MA.