Screening for Problem Drinking

Comparison of CAGE and AUDIT

Katharine A. Bradley, MD, MPH, Kristen R. Bush, MPH, Mary B. McDonnell, MS, Timothy Malone, MS, Stephan D. Fihn, MD, MPH, for the Ambulatory Care Quality Improvement Project (ACQUIP)

OBJECTIVE: To compare self-administered versions of three questionnaires for detecting heavy and problem drinking: the CAGE, the Alcohol Use Disorders Identification Test (AUDIT), and an augmented version of the CAGE.

DESIGN: Cross-sectional surveys.

SETTING: Three Department of Veterans Affairs general medical clinics.

PATIENTS: Random sample of consenting male outpatients who consumed at least 5 drinks over the past year ("drinkers"). Heavy drinkers were oversampled.

MEASUREMENTS: An augmented version of the CAGE was included in a questionnaire mailed to all patients. The AUDIT was subsequently mailed to "drinkers." Comparison standards, based on the tri-level World Health Organization alcohol consumption interview and the Diagnostic Interview Schedule, included heavy drinking (>14 drinks per week typically or ≥ 5 drinks per day at least monthly) and active DSM-III-R alcohol abuse or dependence (positive diagnosis and at least one alcohol-related symptom in the past year). Areas under receiver operating characteristic curves (AUROCs) were used to compare screening questionnaires.

MAIN RESULTS: Of 393 eligible patients, 261 (66%) returned the AUDIT and completed interviews. For detection of active alcohol abuse or dependence, the CAGE augmented with three more questions (AUROC 0.871) performed better than either the CAGE alone or AUDIT (AUROCs 0.820 and 0.777, respectively). For identification of heavy-drinking patients, however, the AUDIT performed best (AUROC 0.870). To identify both heavy drinking and active alcohol abuse or dependence, the augmented CAGE and AUDIT both performed well, but the AUDIT was superior (AUROC 0.861).

CONCLUSIONS: For identification of patients with heavy drinking or active alcohol abuse or dependence, the self-administered AUDIT was superior to the CAGE in this population.

KEY WORDS: alcohol abuse; screening; CAGE; AUDIT; alcohol drinking.

J GEN INTERN MED 1998;13:379-388.

Even though 10% to 36% of patients seen by primary care physicians suffer from alcohol abuse or dependence, these disorders often remain unrecognized by primary care providers.¹ Screening for problem drinking using validated questionnaires is therefore recommended.² The 4-item CAGE questionnaire is a brief, effective screening questionnaire for alcohol abuse and dependence.^{1,3-13} However, the CAGE asks about alcohol problems "ever" experienced, so many patients who respond "yes" to the CAGE questions may have had alcohol problems only in the distant past. A more useful questionnaire screen would predominantly identify patients with "active" alcohol abuse or dependence, who meet diagnostic criteria, and have had at least one alcohol-related symptom in the past year.

Increasingly, clinicians are also interested in identifying heavy drinkers, i.e., patients who are at increased risk of adverse consequences of drinking regardless of whether they meet diagnostic criteria for abuse or dependence.¹⁴ Although there is no consensus on the definition of heavy drinking,¹⁵ we define heavy drinking as more than 14 drinks per week, or 5 or more drinks per day at least monthly. Average daily intake of two standard-sized drinks increases the relative risk of alcohol-related problems 6- to 12-fold.¹⁶ Episodic drinking of 5 or more drinks also places individuals at increased risk of alcohol-related problems.^{16,17} Moreover, brief primary care interventions with heavy drinkers decrease their subsequent alcohol consumption, and improve other health-related outcomes.^{18,19} Alcohol screening questionnaires will be most helpful, therefore, if they identify heavy drinking as well as active alcohol abuse or dependence.

The CAGE appears to be a poor screen for heavy drinking,²⁰ and the best method for identification of primary care patients who drink heavily or have active alcohol abuse or dependence has not been established. The Alcohol Use Disorders Identification Test (AUDIT) was specifically designed to identify heavy drinkers,²¹ and explicitly addresses alcohol-related problems and symptoms of dependence over the past year. The only published comparison of the AUDIT and CAGE in a primary care population used lifetime alcohol abuse or dependence as the comparison standard.¹³ Not surprisingly, given the AUDIT's explicit focus on past-year symptoms, it performed poorly compared to the CAGE for identifying alcohol abuse and dependence ever.¹³ Emergency department studies have found the AUDIT to perform better than the CAGE questionnaire as a screen for current alcohol abuse or dependence.^{7,8}

Received from Health Services Research and Development (KAB, MBM, SDF), Addictions Treatment Center (KRB), and Medical Service (KAB, SDF), VA Puget Sound Health Care System (Seattle Division, Wash.); and Departments of Medicine (KAB, SDF), Health Services (SDF), and Biostatistics (TM), University of Washington, Seattle.

Address correspondence and reprint requests to Dr. Bradley: Health Services Research and Development, Mailstop 152, VA Puget Sound Health Care System (Seattle Division), 1660 S. Columbian Way, Seattle, WA 98108.

The length of the 10-item AUDIT, however, is a disadvantage. A shorter questionnaire for identification of heavy drinking and active alcohol-related problems is the combination of questions about recent alcohol consumption from the CAGE screen and a question about the patient's own assessment of a drinking problem.^{22,23} We refer to this 7-question approach as an "augmented CAGE" (Appendix A). To our knowledge, the AUDIT has never been directly compared to an augmented CAGE.

We had the opportunity to evaluate self-administered CAGE, AUDIT, and augmented CAGE questionnaires in conjunction with a Veterans Affairs (VA) Ambulatory Care Quality Improvement Project. The purpose of this report is to compare the performance of the CAGE, AUDIT, and augmented CAGE questionnaires with interview-based comparison standards for heavy drinking and active alcohol abuse or dependence.

METHODS

Subjects

We surveyed general medical outpatients at three VA Medical Centers in 1994 as part of a longitudinal study of continuous quality improvement in general medicine clinics. All demographic information on patients in this study was obtained from the VA Decentralized Hospital Computing Program.

All patients ($N = 9,513$) who had seen a general medical provider within 1 year were sent the Health History Questionnaire. Of these, 330 patients were excluded because questionnaires were undeliverable or patients indicated they did not wish to participate (Fig. 1). Of 9,183 eligible patients, 6,116 (67%) returned the questionnaire. All 2,875 patients who had had 5 or more drinks over the past year were considered "drinkers" and mailed the Drinking Practices Questionnaire.

For this study, we selected a random sample of 447 patients who were mailed the Drinking Practices Questionnaire, and who indicated on the Health History Questionnaire that they were "willing to receive one or two telephone calls regarding (their) . . . health and health-related habits." Eighty-five percent of responding patients answered this question "yes." Sampling was stratified with heavy drinkers oversampled 2:1 in order to test the performance of the screening questionnaires in adequate numbers of heavy drinkers. Patients who reported on the Health History Questionnaire that they drank 5 or more drinks per typical day or averaged more than 14 drinks per typical week were considered heavy drinkers. Only six women met inclusion criteria and were excluded from analyses.

Questionnaire Measures

In the Health History Questionnaire, alcohol-related questions followed items on smoking and preceded items

FIGURE 1. Study populations.

on dietary practices. Four CAGE questions and the question "Have you ever had a drinking problem?" were followed by AUDIT questions 1 and 2 about typical alcohol consumption (Appendix A). We scored the CAGE in the traditional manner with each "yes" response assigned 1 point. The "augmented CAGE" was scored as the sum of the CAGE score, plus 1 point if a patient indicated any history of a drinking problem, and 1 point if he reported heavy drinking.

The Drinking Practices Questionnaire included the AUDIT, a retrospective drinking diary, and questions about recent changes in drinking, motivation to change, and provider advice about drinking, and was mailed within 4 weeks of return of the Health History Questionnaire. The AUDIT was scored in the standard manner, with possible scores ranging from 0 to 40.²⁴ The AUDIT and the two versions of the CAGE were scored if more than half the items in the screening instrument were answered, with missing items assigned 0 points.

Interview Measures

The telephone interview included verbal consent, followed by a modified version of the tri-level World Health Organization (WHO) interview on alcohol consumption that asked patients to give detailed descriptions of the types and amounts of alcoholic beverages that they drank

on low-, medium-, and high-level drinking days.^{21,25,26} It also asked patients for the number of low-, medium-, and high-level drinking days in the past month and in a typical month of the past year. Reported consumption was converted to standard drink units using recommended conversion standards.²⁷ After questions about alcohol consumption, interviewers administered the alcohol module of a computerized version of the Diagnostic Interview Schedule (C-DISR, Interview Manager, version 2.0).²⁸ This standardized instrument, used in previous studies of alcohol screening questionnaires,^{1,4,6,29,30} provides a lifetime diagnosis of DSM-III-R alcohol abuse or dependence, as well as the time of the most recent alcohol-related symptom and the number of symptoms in the past year. The entire interview usually lasted 10 to 20 minutes (range, 5 to 40 minutes). Half of the eligible patients were called for interviews before the Drinking Practices Questionnaire was mailed, and half were telephoned after it was returned, to evaluate whether a preceding interview about drinking habits affected reported consumption on the Drinking Practices Questionnaire.

Telephone interviews were completed by five experienced interviewers who were blind to all results of questionnaire screening. To ensure comparable styles of interview administration and coding, interviews were monitored periodically, and all forms were reviewed for completeness and consistency by one of the authors (KAB). Direct assessment of interrater reliability was not possible because the already heavy questionnaire burden on patients in this study precluded performing repeated interviews within a short time period.

Interview Comparison Standards

For analyses, patients were classified as heavy drinkers if they drank more than 14 drinks per week in a typical month, or if they drank 5 or more standard drinks in a day at least monthly, according to interviews.^{15,17} Patients were classified as having "lifetime alcohol abuse or dependence" if they met DSM-III-R criteria for alcohol abuse or dependence at any time in their lives. A subset of patients with lifetime alcohol abuse or dependence who had had one or more symptoms in the past year were classified as having "active alcohol abuse or dependence." A fourth category included patients with heavy drinking or active alcohol abuse or dependence in the past year.

Analyses

Sensitivities, specificities, and positive likelihood ratios for questionnaire screens compared with interview comparison standards were calculated in the standard manner.³¹ We used receiver operating characteristic (ROC) curves to compare the CAGE, the augmented CAGE, and the AUDIT. Receiver operating characteristic curves plot sensitivity versus (1 – specificity), and the area under an ROC curve (AUROC) measures the overall performance of

a screening questionnaire, allowing comparison of two or more screening instruments. Receiver operating characteristic curves were constructed using scores of 1 to 4 for the CAGE, 1 to 6 for the augmented CAGE, and 4 to 17 for the AUDIT. The area under each curve was calculated nonparametrically.^{32,33} To test the hypothesis that the areas under two ROC curves were significantly different, we used the *z* statistic, adjusting for the correlation between AUROCs when derived from the same population.³⁴

We considered several approaches to dealing with missing data, and we present analyses based on all patients who returned Drinking Practices Questionnaires and completed interviews. We were concerned that excluding respondents without complete data for all three questionnaire measures would overestimate sensitivity. For example, using an extreme case, a screening questionnaire that was left blank by 50% of problem drinkers could identify all problem drinkers who completed the questionnaire (100% sensitivity based on completers), but in practice would only identify 50% of problem drinking patients in the entire screened population. For this reason the most conservative estimates of sensitivity are based on the entire study population, and in our main analyses respondents missing a score for the CAGE, augmented CAGE, or AUDIT were assigned a score of 0 for the missing instrument. To determine how much these assumptions about missing responses affected our results, we repeated analyses using only interviewed patients who completed all items in all three screening instruments. In addition, we excluded only those patients who were missing a score for each screening questionnaire, and repeated analyses pertaining to that specific questionnaire.

To evaluate nonresponse bias, we used the Wilcoxon Rank-Sum Test to compare responses to Health History Questionnaire alcohol questions among Drinking Practices Questionnaire respondents and nonrespondents. We hypothesized that patients who drank more or had more alcohol-related problems would be more likely to be nonrespondents. To evaluate the effect of embedding the CAGE in the Health History Questionnaire, while the AUDIT was part of a questionnaire explicitly addressing drinking practices (context response bias), we compared responses to AUDIT questions 1 and 2 included in both questionnaires, using Wilcoxon Signed-Rank Test for matched pairs. We hypothesized that patients would report higher consumption on questions embedded in a general health history questionnaire.

Statistics were performed using SPSS and Excel. S-Plus, version 3.3, and a program written by one of the authors (TM) were used for graphical depiction of ROCs and calculation of the AUROCs. All *p* values are two-tailed.

RESULTS

Of 441 male patients who returned Health History Questionnaires and were selected for interviews, 48 (10.9%) were ineligible due to being too ill or deaf to be in-

interviewed ($n = 5$), having no telephone ($n = 24$), or not answering repeated telephone calls over a 3-week period ($n = 19$). Of the remaining 393 patients, 110 (28.0%) did not return Drinking Practices Questionnaires, and 22 (5.6%) refused interviews. A total of 261 patients (66.4% of eligible) completed interviews and returned Drinking Practices Questionnaires (Fig. 1).

Table 1 describes demographic and screening characteristics of the entire population of Health History Questionnaire respondents, drinkers mailed the Drinking Practices Questionnaire and respondents to that questionnaire, and participants and nonparticipants in the interview study. Unfortunately, VA Decentralized Hospital Computing Program data on ethnicity were frequently missing. The underlying clinical populations were, however, predominantly white (89.9% of the 65% with ethnicity reported). Answers to a question at the end of the Health History

Questionnaire revealed that about 80% of questionnaires were completed by patients themselves; 12% were filled out by patients and their spouses or partners; and 6% were filled out by patients' spouses or someone else.

Of 261 study participants, 127 (49%) met interview criteria for lifetime alcohol abuse or dependence, 56 (22%) met criteria for active alcohol abuse and dependence, 89 (34%) met criteria for heavy drinking, and 105 (40%) met our criteria for "heavy drinking and/or active alcohol abuse or dependence." Among the 261 study participants, 39% indicated on the AUDIT having had 6 or more drinks in a day in the past year (Table 1), a much higher proportion than reported more than 14 drinks in a typical week.

Comparison of AUROCs (Table 2) shows that for identification of active or lifetime alcohol abuse or dependence (Figs. 2a and 2b), the augmented CAGE performed better than the CAGE alone or the AUDIT ($p < .0001$). For detec-

Table 1. Characteristics of Health History Questionnaire (HHQ) Respondents, Drinkers Who Were Mailed and Responded to the Drinking Practices Questionnaire (DPQ), and Drinkers Randomly Selected for the Interview Study*

Characteristic	All HHQ Respondents, % ($n = 6,116$)	Drinkers [†] Based on HHQ		Drinkers [†] Randomly Selected for the Interview Study	
		Total, % ($n = 2,875$)	Respondents, % ($n = 1,911$)	Participants, % ($n = 261$)	Nonparticipants [‡] , (%) ($n = 186$)
Demographics					
Male gender	96.6	97.1	97.5	100	96.7
Age in years					
<50	13.1	15.2	12.5	7.7	12.4
50–59	13.4	14.4	13.0	6.9	19.9
60–69	35.6	35.8	37.2	36.2	37.6
≥70	37.1	33.7	36.7	49.2	30.1
Site					
Boston, Mass.	19.1	19.8	16.5	25.7	38.2
Seattle, Wash.	51.1	46.7	49.0	38.3	30.6
White River Junction, Vt.	29.8	33.5	34.4	36.0	31.2
HHQ responses					
CAGE score ≥2	22.7	25.5	23.2	33.0	36.6
"Ever had a drinking problem"	25.3	24.7	22.6	31.4	38.7
≥14 drinks per week [§]	8.3	16.9	15.0	20.7	29.0
>4 drinks per typical drinking day [¶]	9.2	15.0	12.9	16.9	25.8
Augmented CAGE ≥2	28.2	33.3	30.8	44.8	49.5
DPQ responses					
AUDIT score ≥4	—	—	44.8	52.9	—
AUDIT score ≥8	—	—	18.6	24.9	—
AUDIT score ≥11	—	—	12.0	16.9	—
≥14 drinks per week [§]	—	—	13.3	16.9	—
>4 drinks per typical drinking day [¶]	—	—	12.7	19.2	—
≥6 drinks in a day, past year [#]	—	—	35.1	39.1	—

*Heavy drinkers were oversampled 2:1 for the interview study; therefore, no statistical comparison is made between the entire population of Health History Questionnaire respondents and the study population; significant differences between these two populations are expected.

[†]Patients who had ≥5 drinks over the past year.

[‡]Patients in this group failed to return Drinking Practices Questionnaires ($n = 110$), were excluded because of illness or unavailability ($n = 48$), or refused to be interviewed ($n = 22$) (see Fig. 1).

[§]Based on mean values of AUDIT questions 1 and 2.

^{||}Significant differences were observed between participants ($n = 261$) and nonparticipants ($n = 186$), when comparing drinks per week, typical daily consumption, and augmented CAGE score, using the Wilcoxon Rank-Sum Test ($p < .05$).

[¶]AUDIT question 2.

[#]AUDIT question 3.

Table 2. Areas Under Receiver Operating Characteristic Curves (AUROCs)

Screen	AUROCs (SE) by Interview Comparison Standards*			
	Active Alcohol Abuse or Dependence (n = 56)	Lifetime Alcohol Abuse or Dependence (n = 127)	Heavy Drinking (n = 89)	Heavy Drinking and/or Active Alcohol Abuse or Dependence (n = 105)
CAGE	0.820 (0.018)	0.722 (0.016)	0.668 (0.018)	0.717 (0.017)
Augmented CAGE	0.871 (0.013)	0.757 (0.012)	0.723 (0.014)	0.769 (0.013)
AUDIT	0.777 (0.011)	0.654 (0.0094)	0.870 (0.0075)	0.861 (0.0069)

*See text for definitions.

tion of heavy drinking (Fig. 2c), however, the AUDIT was a superior screening test to the augmented CAGE ($p < .0001$), which in turn performed significantly better than the CAGE alone ($p < .0001$). For identifying heavy drinking and active alcohol abuse or dependence (Fig. 2d), the AUDIT performed better than either the CAGE ($p < .0001$) or the augmented CAGE ($p < .0001$). Again the augmented CAGE performed significantly better than the CAGE alone ($p < .0001$).

All screening questionnaires were more sensitive for active than for lifetime alcohol abuse or dependence (Table 3). However, using traditional cutpoints, the CAGE (≥ 2) missed 47% of patients with heavy drinking or active alcohol abuse or dependence, and the AUDIT (≥ 8) missed 45%; the augmented CAGE (≥ 2) only missed 28%. Using lower cutpoints, however, the CAGE and augmented CAGE (≥ 1) had sensitivities of 77% and 87%, respectively, and the AUDIT (≥ 4) also had a sensitivity of 87%.

The AUDIT had higher specificity for heavy drinking or active alcohol abuse or dependence (Table 3). Using traditional cutpoints, patients without heavy drinking or active alcohol abuse or dependence in the past year, according to the interviews, were much less likely to screen positive on the AUDIT (4%), than on the CAGE or augmented CAGE (19% or 25%, respectively). This advantage is reflected in the AUDIT's higher positive likelihood ratios. Even with a CAGE or augmented CAGE score of 4 or more, the positive likelihood ratios for these questionnaires were lower than for an AUDIT score of 6 or more.

Response Biases

Nonrespondents to the Drinking Practices Questionnaire ($n = 964$) were significantly more likely than respondents ($n = 1,911$) to be heavy drinkers and problem drinkers as determined by the Health History Questionnaire alcohol consumption questions and CAGE scores (data not shown). Similarly, nonparticipants in the interview study ($n = 186$) reported drinking significantly more and had higher scores on the augmented CAGE than participants ($n = 261$; data not shown). We also compared responses to two identical consumption questions (AUDIT 1–2) on the Health History Questionnaire and the Drinking Practices Questionnaire, and responses were not significantly different on the two instruments (data not shown). Comparing patients interviewed before or after

the Drinking Practices Questionnaire, the response rate to the Drinking Practices Questionnaire was significantly higher among patients randomly assigned to be interviewed before that questionnaire. However, there were no significant differences between the randomly assigned groups with regard to consumption or problems reported on the Drinking Practices Questionnaire (data not shown).

Secondary Analyses

To evaluate the effect of our assumptions regarding missing values, we repeated analyses twice using different inclusion criteria. Thirty-five patients included in our main analyses had scores computed based on incomplete CAGE ($n = 8$), augmented CAGE ($n = 8$), or AUDIT questionnaires ($n = 27$), because more than half the questionnaire items were completed. Repeated analyses, including 213 patients who had completed all items of all three screening questionnaires, did not change our main conclusions. Areas under ROC curves ranged from 0.03 higher (AUDIT compared with active alcohol abuse or dependence) to 0.004 lower (CAGE compared with lifetime alcohol abuse or dependence).

Excluding only the 14 patients who were missing more than half of the CAGE questions (5 patients), augmented CAGE questions (5 patients), or AUDIT questions (10 patients), rather than assigning them a score of 0, did not change any of our conclusions either (data not shown). Sensitivities ranged from 0 to 3 percentage points higher, and some specificities were lower by less than 1 percentage point.

DISCUSSION

This study directly compared the AUDIT and CAGE questionnaires to interview comparison standards for heavy drinking and active alcohol abuse or dependence in a male primary care population. In addition, we compared the AUDIT to a briefer, previously described, augmentation of the CAGE.^{22,23,35} All screening questionnaires were evaluated as self-administered instruments, with the two versions of the CAGE embedded in a Health History Questionnaire, and the AUDIT included in a questionnaire explicitly addressing drinking practices. Not surprisingly, given the past-year time frame of the AUDIT,¹³ the CAGE

FIGURE 2. Comparison of AUROCs for identification of active or lifetime alcohol abuse or dependence: (a) Active alcohol abuse and dependence; (b) lifetime alcohol abuse and dependence; (c) heavy drinking; (d) heavy drinking and/or active alcohol abuse or dependence. See text for definitions of comparison standards and the augmented CAGE; numbers in parentheses are standard errors.

and its augmented version performed better than the AUDIT for identification of lifetime alcohol abuse or dependence. Interestingly, however, the augmented version of the CAGE performed significantly better than the AUDIT for detection of active alcohol abuse or dependence. For identification of patients with either heavy drinking, and/or active alcohol abuse or dependence in the past year, however, the AUDIT performed significantly better than the CAGE or augmented CAGE.

Sensitivities of the CAGE and AUDIT for alcohol abuse or dependence in the present study are lower than those reported in many previous studies (Table 4). We suspect several factors contributed to the lower sensitivities in the present study. We surveyed a predominantly white population, and screening questionnaires appear to be less sensitive in white than African-American populations.^{7,8} In addition, previous studies had one interviewer administer both screening questionnaires and compari-

Table 3. Sensitivity, Specificity, and Positive Likelihood Ratios for Questionnaire Screens Compared with Each Interview Comparison Standard

Screen	Positive Screen	Comparison Standards*							
		Active Alcohol Abuse or Dependence (n = 56)		Lifetime Alcohol Abuse or Dependence (n = 127)		Heavy Drinking (n = 89)		Heavy Drinking and/or Active Alcohol Abuse or Dependence (n = 105)	
		Sensitivity/Specificity, %	+ LR†	Sensitivity/Specificity, %	+ LR	Sensitivity/Specificity, %	+ LR	Sensitivity/Specificity, %	+ LR
CAGE	≥1	89/53	1.9	72/59	1.8	74/53	1.6	77/58	1.9
	≥2	77/79	3.7	53/86	3.7	49/75	2.0	53/81	2.8
	≥3	50/93	7.3	30/97	10.0	29/91	3.1	29/92	3.7
Augmented CAGE	≥1	98/46	1.8	82/54	1.8	85/49	1.7	87/53	1.8
	≥2	91/68	2.8	65/74	2.5	70/68	2.2	72/74	2.8
	≥3	73/83	4.4	47/89	4.2	44/79	2.1	49/85	3.2
AUDIT	≥4	55/93	7.6	33/97	11.1	35/91	3.9	34/94	5.4
	≥4	79/54	1.7	61/54	1.3	92/67	2.8	87/70	2.9
	≥5	75/69	2.4	56/74	2.1	82/81	4.3	77/84	4.8
	≥6	70/78	3.2	49/84	3.0	70/87	5.4	66/90	6.8
	≥7	66/83	4.0	43/88	3.6	62/91	6.5	59/94	10.2
	≥8	66/86	4.8	39/89	3.5	57/92	6.9	55/96	12.3
	≥9	61/89	5.4	36/92	4.4	53/94	8.9	50/97	19.7

*See text for definitions; total n = 261.

† + LR indicates positive likelihood ratio (Sensitivity/1 - Specificity).

son standards at the same sitting, potentially biasing results toward higher screening test performance owing to response consistency bias.³⁸ The DSM-III-R criteria that we used may be more inclusive than earlier DSM-III criteria.³⁹

In addition, the lower sensitivities in our study may reflect the fact that we evaluated self-administered ques-

tionnaires. Most previous studies of the AUDIT and CAGE have evaluated these instruments as interviewer-administered questionnaires.^{1,4-13,30,36} Three studies of interviewer-administered CAGE questionnaires in elderly VA populations similar to ours found the sensitivity of the CAGE (≥2) for DSM-III-R lifetime alcohol abuse or depen-

Table 4. Previous Studies Comparing the CAGE and AUDIT with Interview Comparison Standards for DSM-III or DSM-III-R Alcohol Abuse or Dependence

Screen	Study*	African American, %	Male, %	Age, Years	Sensitivity/Specificity	
					Past Year, %	Lifetime, %
CAGE ≥ 2	Primary care ^{1,4}	80	30†	NA†	—	74/91†
			100†	>60†		69/84†
	Primary care ^{5,6}	57	66	mean 41	100/68	91/84
	ED noninjured ^{7,8}	0†	100	85% < 50	58/90†	—
		100†			84/80†	
	Primary care ⁹	25	40	>65	—	48/99
	VA nursing home ¹⁰	NA	97	>50	—	82/90
	VA urgent care ¹¹	NA	100	mean 51	—	80/96
	VA primary care ¹²	37	100	mean 54	100/61	78/76
	VA primary care ¹³	NA	100	≥65	—	63/82
AUDIT ≥ 8	Primary care ³⁶	Predominantly	48	Mean 45	96/96	63/96
	Primary care ³⁰	4	40	Mean 43	—	38/95
	VA primary care ¹³	NA	100	≥65	—	33/91
	ED noninjured ^{7,8}	0†	100	85% < 50	91/86†	—
		100†			90/75†	
	Primary care ³⁷	NA	About 1/3	NA	51/96	—

*Studies of men in clinical settings that used interview-based diagnoses of DSM or ICD alcohol abuse or dependence for comparison standards are included; all ≥18 years old unless noted; ED indicates emergency department; NA, information not available.

† Analyses stratified by these demographic characteristics.

dence to be 78% to 82%.^{10–12} Our sensitivity for the self-administered CAGE (≥ 2) compared with the same comparison standard was 53% (95% confidence interval [CI] 44%, 61%). Similarly, the AUDIT (≥ 8) had a lower sensitivity for active DSM-III-R alcohol abuse or dependence in our population (66%; 95% CI 54%, 78%) than in a study in which the AUDIT was interviewer-administered (96%).³⁶ Previous studies comparing self-administered CAGE and AUDIT questionnaires with self-administered comparison standards have reported lower sensitivities, closer to ours: 39% for the CAGE (≥ 2) screening for lifetime alcohol abuse or dependence³; and 61% for the AUDIT (≥ 8) screening for current alcohol abuse or dependence.⁴⁰

No previous primary care study has compared a self-administered CAGE, AUDIT, or augmented CAGE questionnaire with an interview comparison standard for heavy drinking. One study compared a self-administered CAGE using a modified 3-month time frame with self-administered questions about alcohol consumption in the same survey, and found a sensitivity of only 14% for heavy drinking.²⁰ A study of army recruits found the CAGE to have a sensitivity of 41% for heavy drinking among men, defined as more than 21 drinks a week.⁴¹

Limitations

Several limitations of this study must be kept in mind while interpreting our results. There is no true “gold standard” for alcohol abuse and dependence or heavy drinking. Apparent misclassification of patients by questionnaire screens may therefore reflect misclassification of some patients by interviews. Patients may have been misclassified because of social desirability bias owing to lack of privacy at home while completing questionnaires or interviews. We were also unable to evaluate interrater reliability for interviews, so that unmeasured interviewer biases could have resulted in misclassification of patients by interviewers. Proxy respondents to questionnaires may have further decreased the internal validity of our results. Finally, we embedded the augmented CAGE in the general Health History Questionnaire, whereas the AUDIT was included in a questionnaire explicitly addressing drinking practices, potentially resulting in unmeasured context response biases.

We studied self-administered screening questionnaires in predominantly older, white, male veterans who drank 5 or more drinks over the past year, limiting the generalizability of our findings to other populations or situations in which alcohol screening questionnaires are administered in face-to-face interviews. We also found evidence of response bias, with heavy and problem drinkers being less likely to complete Drinking Practices Questionnaires or interviews. Our data therefore do not predict the performance of alcohol screening questionnaires in the heavy drinkers who did not complete the Drinking Practices Questionnaire or interview. However, nonresponse

bias most likely exists in the majority of clinical and research alcohol screening programs⁴² and we believe it was a strength of this study that we were able to measure drinking-related response biases.

Clinical Implications

Our findings confirm that the choice of a screening instrument should depend on the goal of screening. The augmented CAGE always performed better than the CAGE alone, and appears to be the optimal questionnaire for identifying patients with active alcohol abuse or dependence who might benefit from referral to addiction specialists. If the purpose of screening is to identify patients with heavy drinking so that clinicians can perform brief interventions aimed at changing drinking practices, the AUDIT had a slight benefit over the augmented CAGE in our population. We suspect that the AUDIT's superior performance for detecting heavy drinking predominantly relates to the third question about the frequency of heavy drinking, as the other two AUDIT consumption questions were also included in the augmented CAGE. Consideration should therefore be given to adding the AUDIT question about frequency of heavy drinking to the augmented CAGE.

Because the goal of alcohol-related screening is to identify high-risk patients and initiate further assessment of alcohol-related problems, sensitivity is more important than specificity in this setting. Self-administered screening questionnaires can identify more than 80% of patients with heavy drinking, or active alcohol abuse or dependence, using lower cutpoints than have been traditionally recommended. One or more points on the CAGE or augmented CAGE, or 4 or more points on the AUDIT, should therefore be considered a positive screen and lead to further evaluation.

Our findings suggest that alcohol-screening questionnaires are effective when embedded in mailed general health surveys. Self-administered questionnaires are probably more practical than interviews for alcohol-related screening, as clinicians are unlikely to remember 7 to 10 items for integration into clinical interviews. Moreover, with improvements in computer-scanning technology, alcohol-related screening may be increasingly included in population-based, mailed surveys to primary care patients. Given recommendations that patients be screened for many health-related behaviors and conditions ranging from smoking and unsafe sexual practices to domestic violence and depression, it is often impractical for all screening to occur during patient visits. In addition, managed care organizations want population-based data regarding patients' high-risk behaviors.⁴³ Although patients who do not respond to such mailed surveys must be screened during clinical encounters, the same computer-scannable survey could be administered by clerical staff. Taking screening off the shoulders of primary care clinicians will not only increase the number of problem drinking patients identified,⁴⁴ but also increase counseling and referral.^{45,46}

This research was supported by Department of Veteran Affairs, Hines Center for Cooperative Studies in Health Services Research, grant 91-007, and Health Services Research and Development, grant SDR 96-002, Ambulatory Care Quality Improvement Project (ACQUIP); a grant from the University of Washington Alcohol and Drug Abuse Institute; and the Health Services Research and Development Field Program and Medical Service, Seattle Division, VA Puget Sound Health Care System.

The authors thank Daniel Kivlahan, PhD, and Daniel Lessler, MD, for their comments on an earlier version of this manuscript.

REFERENCES

- Buchsbaum DG, Buchanan RG, Centor RM, Schnoll SH, Lawton MJ. Screening for alcohol abuse using CAGE scores and likelihood ratios. *Ann Intern Med.* 1991;115:774-7.
- Connors GJ. Screening for alcohol problems. In: Allen JP, Columbus M, eds. *Assessing Alcohol Problems: A Guide for Clinicians and Researchers.* Bethesda, Md: National Institutes of Health; 1995:17-29. NIH publication 95-3745.
- Barry KL, Fleming MF. Computerized administration of alcoholism screening tests in a primary care setting. *J Am Board Fam Pract.* 1990;3:93-8.
- Buchsbaum DG, Buchanan RG, Welsh J, Centor RM, Schnoll SH. Screening for drinking disorders in the elderly using the CAGE questionnaire. *J Am Geriatr Soc.* 1992;40:662-5.
- Chan AK, Pristach EA, Welte JW. Detection by the CAGE of alcoholism or heavy drinking in primary care outpatients and the general population. *J Subst Abuse.* 1994;6(2):123-35.
- Chan AK, Pristach EA, Welte JW, Russell M. Use of the TWEAK test in screening for alcoholism/heavy drinking in three populations. *Alcohol Clin Exp Res.* 1993;17(6):1188-92.
- Cherpitel CJ. Analysis of cut points for screening instruments for alcohol problems in the emergency room. *J Stud Alcohol.* 1995;56:695-700.
- Cherpitel CJ, Clark WB. Ethnic differences in performance of screening instruments for identifying harmful drinking and alcohol dependence in the emergency room. *Alcohol Clin Exp Res.* 1995;19(3):628-34.
- Jones TV, Lindsey BA, Yount P, Soltys R, Farani-Enayat B. Alcoholism screening questionnaires: are they valid in elderly medical outpatients? *J Gen Intern Med.* 1993;8:674-8.
- Joseph CL, Ganzini L, Atkinson RM. Screening for alcohol use disorders in the nursing home. *J Am Geriatr Soc.* 1995;43:368-73.
- Liskow B, Campbell J, Nickel EJ, Powell BJ. Validity of the CAGE questionnaire in screening for alcohol dependence in a walk-in (triage) clinic. *J Stud Alcohol.* 1995;56:277-81.
- Magruder-Habib K, Stevens HA, Alling WC. Relative performance of the MAST, VAST, and CAGE versus DSM-III-R criteria for alcohol dependence. *J Clin Epidemiol.* 1993;46(5):435-41.
- Morton J, Jones T, Mangano M. Performance of alcoholism screening questionnaires in elderly veterans. *Am J Med.* 1996;101:153-9.
- Bradley KA, Donovan DM, Larson EB. How much is too much? Advising patients about safe levels of alcohol consumption. *Arch Intern Med.* 1993;153:2734-40.
- Sanchez-Craig M, Wilkinson A, Davila R. Empirically based guidelines for moderate drinking: one year results from three studies with problem drinkers. *Am J Public Health.* 1995;85(6):823-8.
- Dawson DA, Grant BF, Harford TC. Variation in the association of alcohol consumption with five DSM-IV alcohol problem domains. *Alcohol Clin Exp Res.* 1995;19(1):66-74.
- Wechsler H, Davenport A, Dowdall G, Moeykens B, Castillo S. Health and behavioral consequences of binge drinking in college. *JAMA.* 1994;272:1672-7.
- Fleming MF, Barry KL, Manwell LB, Johnson K, London R. Brief physician advice for problem alcohol drinkers. *JAMA.* 1997;277(13):1039-45.
- Bien TH, Miller WR, Tonigan S. Brief interventions for alcohol problems: a review. *Addiction.* 1993;88:315-36.
- Adams WL, Barry KL, Fleming MF. Screening for problem drinking in older primary care patients. *JAMA.* 1996;276:1964-7.
- 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.
- Fleming MF, Barry KL. A three-sample test of a masked alcohol screening questionnaire. *Alcohol Alcohol.* 1991;26(1):81-91.
- Wallace P, Haines A. Use of a questionnaire in general practice to increase the recognition of patients with excessive alcohol consumption. *BMJ.* 1985;290:1949-52.
- Babor TF, Grant M. From clinical research to secondary prevention—international collaboration in the development of the Alcohol Use Disorders Identification Test (AUDIT). *Alcohol Health Res World.* 1989;13(4):371-4.
- Saunders JB, Aasland OG, Amundsen A, Grant M. Alcohol consumption and related problems among primary health care patients: WHO collaborative project on early detection of harmful alcohol consumption, I. *Addiction.* 1993;88:349-62.
- Saunders JB, Aasland OG. WHO Collaborative Project on Identification and Treatment of Persons with Harmful Alcohol Consumption: Report on Phase I, Development of a Screening Instrument. Geneva, Switzerland: World Health Organization, Division of Mental Health; 1987.
- Miller WR, Heather N, Hall W. Calculating standard drink units: international comparisons. *Br J Addict.* 1991;86:43-7.
- Robins LN, Helzer JE, Croughan J, Ratcliff KS. National Institute of Mental Health Diagnostic Interview Schedule. *Arch Gen Psychiatry.* 1981;38:381-9.
- Fleming MF, Barry KL, MacDonald R. The Alcohol Use Disorders Identification Test (AUDIT) in a college sample. *Int J Addict.* 1991;26(11):1173-85.
- Schmidt A, Barry K, Fleming MF. Detection of problem drinkers: the Alcohol Use Disorders Identification Test. *South Med J.* 1995;88(1):52-9.
- Ingelfinger JA, Mosteller F, Thibodeau LA, Ware JH. *Biostatistics in Clinical Medicine.* 3rd ed. New York, NY: McGraw-Hill; 1994.
- Zweig MH, Campbell G. ROC plots: a fundamental evaluation tool in clinical medicine. *Clin Chem.* 1993;39:561-77.
- Centor RM, Schwartz JS. An evaluation of methods for estimating the area under the ROC curve. *Med Decis Making.* 1985;5:149-56.
- Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology.* 1983;148:839-43.
- Cutler SF, Wallace PG, Haines AP. Assessing alcohol consumption in general practice patients—a comparison between questionnaire and interview. *Alcohol Alcohol.* 1988;23(6):441-50.
- Isaacson HJ, Butler R, Zacharek M, Tzelepis A. Screening with the Alcohol Use Disorders Identification Test (AUDIT) in an inner city population. *J Gen Intern Med.* 1994;9:550-3.
- Volk RJ, Steinbauer JR, Cantor SB, Holzer CE III. The Alcohol Use Disorders Identification Test (AUDIT) as a screen for at-risk patients of different racial/ethnic backgrounds. *Addiction.* 1997;92(2):197-206.
- Harford TC. The effects of order of questions on reported alcohol consumption. *Addiction.* 1994;89:421-4.
- Schuckit MA, Hesselbrock V, Tipp J, Anthenelli R, Bucholz K, Radziminski S. A comparison of DSM-III-R, DSM-IV and ICD-10 substance use disorders diagnoses in 1922 men and women subjects in the COGA study. Collaborative Study on the Genetics of Alcoholism. *Addiction.* 1994;89(12):1629-38.

40. Barry KL, Fleming MF. The Alcohol Use Disorders Identification Test (AUDIT) and the SMAST-13: predictive validity in a rural primary care sample. *Alcohol*. 1993;28(1):33-42.
41. Fertig JB, Allen JP, Cross GM. CAGE as a predictor of hazardous alcohol consumption in U.S. Army personnel. *Alcohol Clin Exp Res*. 1993;17(6):1184-7.
42. Richmond R, Heather N, Wodak A, Kehoe L, Webster I. Controlled evaluation of a general practice based brief intervention for excessive drinking. *Addiction*. 1995;90:119-32.
43. CDC. Health risk factor surveys of commercial plan- and Medicaid-enrolled members of health-maintenance organizations—Michigan 1995. *MMWR*. 1997;46(39):923-6.
44. Buchsbaum DG, Buchanan RG, Poses RM, Schnoll SH, Lawton MJ. Physician detection of drinking problems in patients attending a general medicine practice. *J Gen Intern Med*. 1992;7(5):517-21.
45. Buchsbaum DG, Buchanan RG, Lawton MJ, Elswick RKJ, Schnoll SH. A program of screening and prompting improves short-term physician counseling of dependent and nondependent harmful drinkers. *Arch Intern Med*. 1993;153(13):1573-7.
46. Goldberg HI, Mullen M, Ries RK, Psaty BM, Ruch BP. Alcohol counseling in a general medicine clinic: a randomized controlled trial to improve referral and show rates. *Med Care*. 1991;29(7 suppl):JS49-56.

APPENDIX A

Screening Instruments

CAGE

1. Have you ever felt you ought to cut down on your drinking?
2. Have people annoyed you by criticizing your drinking?
3. Have you ever felt bad or guilty about your drinking?
4. Have you ever had a drink first thing in the morning to steady your nerves or get rid of a hangover?

Alcohol Use Disorders Identification Test (AUDIT)

1. How often did you have a drink containing alcohol in the past year?
2. How many drinks containing alcohol did you have on a typical day when you were drinking in the past year?
3. How often did you have 6 or more drinks on one occasion in the past year?
4. How often during the last year have you found that you were not able to stop drinking once you had started?
5. How often during the last year have you failed to do what was normally expected from you because of drinking?
6. How often during the last year have you needed a first drink in the morning to get yourself going after a heavy drinking session?
7. How often during the last year have you had a feeling of guilt or remorse after drinking?
8. How often during the last year have you been unable to remember what happened the night before because you had been drinking?
9. Have you or someone else been injured as a result of your drinking?
10. Has a relative or friend or doctor or other health care worker been concerned about your drinking or suggested you cut down?

Augmented CAGE

1. CAGE questions
2. AUDIT questions 1 and 2 (above)
3. Have you ever had a drinking problem?