Item Bias in the CAGE Screening Test for Alcohol Use Disorders

Robert J. Volk, PhD, Scott B. Cantor, PhD, Jeffrey R. Steinbauer, MD, Alvah R. Cass, MD

OBJECTIVE: To explore potential item bias in the CAGE questions (mnemonic for cut-down, annoyed, guilty, and eye-opener) when used to screen for alcohol use disorders in primary care patients.

DESIGN AND SETTING: Cross-sectional study, conducted in a university-based, family practice clinic, with the presence of an alcohol use disorder determined by structured diagnostic interview using the Alcohol Use Disorder and Associated Disabilities Interview Schedule.

PATIENTS: A probability sample of 1,333 adult primary care patients, with oversampling of female and minority (African–American and Mexican–American) patients.

MAIN RESULTS: Unadjusted analyses showed marked differences in the sensitivity and specificity of each CAGE question against a lifetime alcohol use disorder, across patient subgroups. Women, Mexican–American patients, and patients with annual incomes above $40,000 were consistently less likely to endorse each CAGE question "yes," after adjusting for the presence of an alcohol use disorder and pattern of alcohol consumption. In results from logistic regression analyses predicting an alcohol use disorder, *cut-down* was the only question retained in models for each of the subgroups. The *guilty* question did not contribute to the prediction of an alcohol use disorder; *annoyed* and *eye-opener* were inconsistent predictors.

CONCLUSIONS: Despite its many advantages, the CAGE questionnaire is an inconsistent indicator of alcohol use disorders when used with male and female primary care patients of varying racial and ethnic backgrounds. Gender and cultural differences in the consequences of drinking and perceptions of problem alcohol use may explain these effects. These biases suggest the CAGE is a poor "rule-out" screening test. Brief and unbiased screens for alcohol use disorders in primary care patients are needed.

KEY WORDS: alcohol use disorders; screening; primary care; sensitivity; specificity.
J GEN INTERN MED 1997;12:763–769.

Although alcohol use disorders are common among primary care patients, physicians are criticized frequently for underrecognition of these problems.[1] Many self-report and biochemical screening tests for alcohol use disorders have been developed to assist physicians in case finding.[2] Nevertheless, the accuracy of alcohol abuse screening tests, when used with male and female patients of different racial and ethnic backgrounds and of varying socioeconomic status, has rarely been considered.[3–5]

The CAGE is a popular self-report measure of alcohol use problems, developed originally to identify the "hidden alcoholic" in hospital settings, where high rates of alcohol disorders are often observed.[6] The CAGE is an acronym for four brief questions: Have you ever felt you should *cut down* on your drinking? Have people *annoyed* you by criticizing your drinking? Have you ever felt bad or *guilty* about your drinking? Have you ever had a drink first thing in the morning (*eye-opener*) to steady your nerves or to get rid of a hangover? The CAGE is considered a covert measure of alcoholism because it does not directly assess alcohol consumption.[7] Its brevity and applicability for varied clinical settings make it particularly attractive as a screen for alcohol use disorders in the primary care setting. This screening tool has been used also in a software package to teach basic concepts of probability revision (Centor R, Allinson J. The ROC Analyzer for Windows; 1995).

The study reported here examines the accuracy of the CAGE in a multiethnic, multiracial sample of primary care patients. The objective of this study is to identify potential bias in the specific CAGE questions when used with male and female patients of varying racial and ethnic backgrounds and, if bias is observed, to explain why the questions perform inconsistently. The term item bias is used to refer to differences in the accuracy of the CAGE questions across patient subgroups. For bias to be present, such differences must be of clinical importance.

METHODS

Setting and Sampling

The study was conducted at the Family Practice Center of The University of Texas Medical Branch at Galveston, a university-based, residency-training clinic serving an ethnically and economically diverse population residing along the Texas Gulf Coast. Data were collected over a 15-month period, from October 1994 to December 1995. The sampling design included oversampling of female patients because of the lower prevalence of alcohol use disorders among women, and oversampling of minority patients from the predominant (African American and Mexican American) groups in the area.

Adult patients scheduled for appointments at the clinic were selected randomly. Patients were contacted by telephone the day before their scheduled office visit and asked to participate in the study, or patients were approached in the clinic waiting room if telephone contact was not possible (approximately 30% of contacts). For patients not keeping their office appointments, the next available patient in the same time slot was approached. This combined sampling strategy resulted in a refusal rate of 5.7%. The final sample included 1,333 patients.

Procedures

Patients completed self-report demographic questionnaires while waiting to see their physicians; after their office visits, they met with project staff to participate in the diagnostic interview. The interview included a diagnostic schedule, the CAGE, and measures of alcohol consumption. All study materials were translated into Spanish by university translation services and back-translated to ensure accuracy. Spanish-speaking interviewers worked with the Mexican American patients (30 patients selected Spanish administration). Interview time was usually less than 45 minutes and varied by the drinking history of the patient.

Measures

Alcohol Use Disorders. Alcohol Use Disorder and Associated Disabilities Interview Schedule (AUDADIS) was used as the diagnostic schedule in this study, to determine the presence of an alcohol use disorder meeting *DSM-IV* diagnostic criteria.[8] The AUDADIS is a structured diagnostic schedule developed for use in the National Longitudinal Alcohol Epidemiologic Survey and designed to be administered by nonclinicians. The test-retest reliability of the AUDADIS for alcohol use disorders is excellent.[9] Current and lifetime alcohol use disorders were considered because some patients with a history of a disorder did not meet criteria for a disorder in the past year.

Alcohol Consumption Patterns. Our measure of alcohol consumption patterns is a quantity/frequency measure, adapted from definitions used by Cahalan et al.[10] and Cherpitel.[11] The following six patterns were identified:

♦ *Abstainers*—never more than 12 drinks of any kind of alcohol in a given year;
♦ *Infrequent drinkers*—did not drink in past 30 days, and did not meet criteria for an abstainer;
♦ *Less frequent, low-quantity drinkers*—drank fewer than 4 days in past 30 days, and drank between 1 and 4 drinks per episode;
♦ *Less frequent, high-quantity drinkers*—drank fewer than 4 days in past 30 days, and drank 5 or more drinks per episode;
♦ *Frequent, low-quantity drinkers*—drank more than 5 days in past 30 days, and drank between 1 and 4 drinks per episode; and
♦ *Frequent, heavy drinkers*—drank more than 5 days in past 30 days, and drank 5 or more drinks per episode.

Data Analysis

For a preliminary analysis, we examined the accuracy of the total CAGE scores (i.e., the sum of "yes" responses to the four questions, giving 1 point for each "yes" response) across patient gender and racial and ethnic subgroups. (In this study, we use the term "accuracy" in a general sense to refer to sensitivity, specificity, and predictive value of a test or question.) Patient educational status and income were included because these two indicators are correlated with race and ethnicity. The percentage of patients screening positive for each subgroup was determined. We calculated likelihood ratios for positive (LR+) and negative (LR−) results, and 95% confidence limits (CL) for the total CAGE score for each patient subgroup, where a cutpoint of 2 defined a positive screen.[12]

For this analysis, the LR+ represents the likelihood that a patient *with* an alcohol use disorder will score 2 or higher (true-positive rate, or sensitivity) divided by the likelihood that a patient *without* an alcohol use disorder will also score 2 or higher (false-positive rate). The larger the LR+, the more accurate the screen in correctly "ruling in" alcohol use disorders. The LR− represents the likelihood that a patient *with* an alcohol use disorder will score below 2 (false-negative rate) divided by the likelihood that a patient *without* an alcohol use disorder will also score less than 2 (true-negative rate, or specificity). The smaller the LR−, the more accurate the screen in correctly "ruling out" alcohol use disorders. The potential for item bias would be suggested if the LRs for the total CAGE scores were inconsistent across patient subgroups.

A series of analyses was used to examine potential item bias in the CAGE questions. We first calculated the percentage of patients answering each question "yes" and then calculated the sensitivity and specificity of each CAGE question against the presence of an alcohol use disorder for each patient subgroup. The LR+, LR−, and associated 95% CL were also calculated. These analyses were unadjusted for potential confounders. Item bias would be suggested if the sensitivity and specificity of the questions varied markedly across patient subgroups, and these differences resulted in poorer accuracy for some subgroups.

We then took a somewhat different approach to the examination of item bias. Logistic regression models were specified in which each CAGE question served as the dependent variable. A hierarchical variable entry approach was used, in which gender and race or ethnicity were entered in the first step and annual income was entered in the second step. In the final step, the presence of an alcohol use disorder (satisfying *DSM-IV* criteria) and the quantity/frequency of alcohol consumption measure were entered.

Item bias would be suggested if gender, or race or ethnicity remained significant, after the alcohol use measures were included in the model. Interaction terms for gender and race were considered also. Odds ratios and 95% CL were reported.

In the final analyses, logistic regression models were specified, in which the CAGE questions served as predictors of an alcohol use disorder. Separate models were estimated for all patients and separately for men, women, whites, African Americans, and Mexican Americans. Item bias would be suggested if different questions were retained in the subgroup models. The CAGE questions were entered in a forward stepwise fashion after controlling for gender, sex, race or ethnicity, and income. These latter analyses yielded the unique (adjusted) effect of each CAGE question in predicting an alcohol use disorder. The analyses were repeated using a backward entry approach, to confirm the previous results.

RESULTS

Demographic characteristics of the sample appear in Table 1. The educational status and income levels were fairly broad, though tending to be lower for minority patients.

Bias in Total CAGE Scores

Table 2 shows the percentage of patients from each subgroup screening positive on the CAGE. Shown also are the LR+ and LR− and the associated 95% CL for the total CAGE score as a screening for alcohol use disorders (a score of 2 or higher on the CAGE is a positive result). The LR+ was higher for women then for men within each racial or ethnic group. The CAGE was most accurate among African–American women and least accurate for African–American men. The clinical importance of the variability in the LR+ estimates can be seen in the predictive value of a positive screen. Assuming a prior probability of an alcohol use disorder of .10, the predictive value of a positive CAGE result would be approximately twice as high for female than for male patients. The LR+ and LR− by educational status and income are presented also in Table 2, with some small differences noted. (As these two variables are highly correlated with race and with each other, we retained only income in subsequent analyses.) Differences across patient subgroups were noted also for the LR− estimates. The lower LR− for African–American women compared with men suggests that the CAGE is a better rule-out test for these patients. These findings are reversed for Mexican–American men and women. Overall, the LR− estimates are higher than would be preferred for a good rule-out test.

Item Bias in Patient Subgroups (Unadjusted Analyses)

Table 3 shows the percentage of patients who answered "yes" to each CAGE question. Overall, men were more likely to answer each question "yes" (related to gender differences in prevalence). For both men and women, Mexican American patients had the lowest rates of "yes" responses.

Table 3 presents also the sensitivity, specificity, and LR+ for each CAGE question. The *cut-down* question was least sensitive for Mexican–American women and more specific for women from each of the three ethnic groups. Fewer than 40% of Mexican–American women with an alcohol use disorder answered this question "yes." Overall, a "yes" response to the *cut-down* question was most accurate (i.e., LR+ is highest) for white and African–American women.

The *annoyed* question was more specific for women, while being poorly sensitive for all groups. A "yes" response was most accurate for women. The *guilty* question was less sensitive for Mexican–American men and women than for patients from the other groups.

Table 1. Sample Characteristics

	Total (*n* = 1,333)	White		African American		Mexican American	
Characteristic		Male (*n* = 165)	Female (*n* = 347)	Male (*n* = 132)	Female (*n* = 339)	Male (*n* = 102)	Female (*n* = 248)
Age, years							
Mean	43	47	45	46	40	45	39
Range	18–86	19–78	18–81	18–77	18–82	19–86	18–77
Education, %							
High school graduate or less	42.9	31.1	31.1	53.8	41.9	56.9	57.3
Some college/vocational training	41.9	32.3	47.8	38.5	49.0	32.4	35.9
College graduate	15.2	36.6	21.0	7.7	9.1	10.8	6.9
Annual income, %							
<$20,000	53.7	28.7	40.4	60.9	71.3	51.0	62.7
$20,000–$39,999	26.5	37.2	29.7	26.6	20.6	25.0	23.8
≥$40,000	19.7	34.1	29.9	12.5	8.1	24.0	13.5

Table 2. Likelihood Ratios and 95% Confidence Limits for CAGE as a Screen for Alcohol Use Disorders: Positive Result Is Two or More Questions Answered "Yes"

Patient Subgroup	Percentage Screening Positive	Likelihood Ratio for a Positive Result*	95% CL	Likelihood Ratio for a Negative Result†	95% CL
Total	15.4	4.85	3.84, 6.08	0.59	0.51, 0.66
White					
Male	30.3	2.58	1.66, 3.93	0.56	0.38, 0.77
Female	13.0	6.26	3.78, 10.14	0.56	0.41, 0.72
African American					
Male	28.0	2.32	1.34, 3.78	0.62	0.39, 0.88
Female	11.2	8.94	5.24, 14.77	0.45	0.28, 0.63
Mexican American					
Male	19.6	3.40	1.59, 7.13	0.67	0.46, 0.88
Female	6.0	9.13	3.54, 23.07	0.76	0.60, 0.89
Education					
High school graduate or less	17.2	4.41	3.17, 6.06	0.55	0.43, 0.68
Some college/vocational training	13.6	5.58	3.78, 8.12	0.59	0.47, 0.71
College graduate	15.3	4.97	2.65, 9.22	0.64	0.48, 0.79
Annual income					
<$10,000	19.0	5.61	3.84, 8.12	0.44	0.32, 0.58
$10,000 to $19,999	13.4	3.02	1.73, 5.07	0.77	0.61, 0.90
$20,000 to $39,999	14.3	3.21	1.93, 5.17	0.73	0.58, 0.87
≥$40,000	14.7	9.12	5.17, 15.96	0.44	0.30, 0.61

Likelihood Ratio Positive is defined as the likelihood of a positive result in patients with an alcohol use disorder, divided by the likelihood of a positive result in patients without a disorder. As the Likelihhod Ratio Positive increases from 1, the probability of an alcohol use disorder also increases with a positive result.

†*Likelihood Ratio Negative is defined as the likelihood of a negative result in patients with an alcohol use disorder, divided by the likelihood of a negative result in patients without a disorder. As the Likelihood Ratio Negative decreases from 1, the probability of an alcohol use disorder also decreases with a negative result.*

The *eye-opener* question was the least frequently endorsed, being most sensitive for African Americans (26%–40%), but poorly sensitive for all groups. Specificity of this question was higher for women and Mexican–American men. A "yes" response was least accurate for white men.

Sociodemographic Predictors of Responses to the CAGE Questions

Table 4 presents results from logistic regression models predicting responses to each of the CAGE questions. (Each CAGE question served as the dependent variable in these logistic regression models.) As Table 4 shows, for each CAGE question, gender of the patient was significant in the final model, with men more likely than women to respond "yes," after controlling for race or ethnicity, income, and the alcohol problems indicators. Race or ethnicity was also significant, as Mexican–American patients were less likely to respond "yes" than were white or African–American patients. Finally, patients with annual incomes less than $10,000 were more likely to endorse each question "yes" compared with patients with annual incomes of $40,000 or more. (Including interaction terms for gender and race did not change significantly the main effects for these variables.)

Item Bias in Patient Subgroups (Adjusted Analyses)

Table 5 presents results from logistic regression models, in which an alcohol use disorder was predicted by the CAGE questions, with adjustments for covariates. Models were estimated for the entire sample (with gender, race or ethnicity, and income serving as covariates) and then separately by gender and race or ethnicity (with the same covariates unless serving as a subgroup selection variable). These effects represent the unique association of each CAGE question with an alcohol use disorder, adjusted for the other CAGE questions and covariates.

Several consistencies are apparent from the analyses reported in Table 5. First, the *cut-down* question remained in each model and had the largest adjusted odds ratio. Second, the *guilty* question was not retained in any of the models considered.

The models also diverged in several areas. For men and white patients, only the *cut-down* question was retained. The *annoyed* question was retained in the model for women but not for models with the other groups. The *eye-opener* question was retained in models for women, African Americans, and Mexican Americans. A reanalysis that used a backward stepwise approach did not change the findings.

**Table 3. Percentage of "Yes" Responses, Sensitivity, Specificity, and Likelihood Ratios as Indicators of
a Lifetime Alcohol Use Disorder, by CAGE Question***

Patient Subgroup	Percentage "Yes" Responses	TPR (Sensitivity)	TNR (Specificity)	Likelihood Ratio for a Positive Result (95% CL)	Likelihood Ratio for a Negative Result (95% CL)
Cut-Down					
Total	23.4	.63	.84	3.95 (3.32, 4.66)	0.44 (0.36, 0.52)
White					
Male	42.4	.76	.69	2.40 (1.74, 3.28)	0.36 (0.20, 0.58)
Female	18.6	.63	.88	5.39 (3.64, 7.82)	0.42 (0.28, 0.58)
African American					
Male	39.4	.68	.67	2.08 (1.37, 2.99)	0.48 (0.25, 0.79)
Female	16.5	.71	.89	6.43 (4.27, 9.36)	0.33 (0.18, 0.52)
Mexican American					
Male	34.3	.67	.77	2.94 (1.78, 4.82)	0.43 (0.24, 0.69)
Female	13.7	.37	.90	3.77 (2.05, 6.62)	0.70 (0.51, 0.86)
Annoyed					
Total	9.4	.24	.93	3.66 (2.64, 5.04)	0.81 (0.74, 0.87)
White					
Male	18.2	.27	.85	1.75 (0.90, 3.27)	0.86 (0.68, 1.02)
Female	7.2	.24	.95	5.11 (2.47, 10.25)	0.80 (0.65, 0.91)
African American					
Male	19.7	.32	.83	1.90 (0.91, 3.68)	0.82 (0.91, 1.02)
Female	6.5	.35	.96	9.94 (4.67, 20.38)	0.67 (0.49, 0.82)
Mexican American					
Male	10.8	.15	.91	1.59 (0.52, 4.61)	0.94 (0.74, 1.08)
Female	4.4	.14	.97	5.07 (1.69, 14.67)	0.88 (0.73, 0.97)
Guilty					
Total	14.4	.37	.90	3.63 (2.82, 4.63)	0.70 (0.62, 0.77)
White					
Male	21.8	.44	.86	3.02 (1.74, 5.18)	0.66 (0.48, 0.83)
Female	13.9	.39	.90	3.90 (2.34, 6.27)	0.68 (0.52, 0.82)
African American					
Male	25.0	.48	.80	2.45 (1.36, 4.16)	0.65 (0.41, 0.89)
Female	12.1	.36	.90	3.64 (1.98, 6.27)	0.71 (0.52, 0.88)
Mexican American					
Male	18.6	.30	.85	2.02 (0.91, 4.32)	0.82 (0.60, 1.02)
Female	5.6	.26	.98	10.95 (4.02, 29.37)	0.76 (0.59, 0.88)
Eye-Opener					
Total	7.6	.21	.95	4.07 (2.82, 5.83)	0.83 (0.77, 0.89)
White					
Male	15.8	.20	.86	1.34 (0.63, 2.75)	0.94 (0.76, 1.09)
Female	5.8	.15	.96	3.50 (1.48, 7.95)	0.89 (0.75, 0.97)
African American					
Male	16.7	.40	.89	3.57 (1.72, 7.09)	0.68 (0.46, 0.88)
Female	5.0	.26	.97	8.83 (3.68, 20.33)	0.76 (0.58, 0.89)
Mexican American					
Male	7.8	.15	.95	2.78 (0.80, 9.44)	0.90 (0.71, 1.02)
Female	3.2	.17	.99	18.26 (4.33, 76.43)	0.84 (0.68, 0.93)

TPR is the true positive rate, or sensitivity; TNR is the true negative rate, or specificity; likelihood ratio of a positive result is ratio of true-positive to false-positive rates; likelihood ratio of a negative result is ratio of false-negative to true-negative rates.

DISCUSSION

This study demonstrates differences in the association between the CAGE questions and alcohol use disorders for male and female patients of varying racial or ethnic backgrounds. These item biases are not explained by differences in prevalence rates of alcohol use disorders or patterns of alcohol use across patient subgroups.

Men were consistently more likely to endorse each CAGE question "yes" after adjustments for race or ethnicity, income, and the alcohol use indicators. This finding may be explained by differences in the distribution of al-

Table 4. Adjusted Odds Ratios (95% Confidence Limits) from Hierarchical Logistic Regression Models, in Which Sociodemographic Factors Serve as Predictors of a "Yes" Response to each CAGE Question*

Variables	Cut-Down	Annoyed	Guilty	Eye-Opener
Sex (0 = female; 1 = male)	2.53[†] (1.85, 3.47)	2.56[†] (1.70, 3.86)	1.77[†] (1.25, 2.51)	2.55[†] (1.63, 3.99)
Race (Mexican American is reference category)				
White	1.69[†] (1.13, 2.53)	2.15[†] (1.22, 3.77)	2.13[†] (1.33, 3.42)	2.19[†] (1.16, 4.13)
African American	1.53[†] (1.02, 2.30)	2.09[†] (1.18, 3.68)	2.24[†] (1.39, 3.61)	2.17[†] (1.14, 4.12)
Annual income ($40,000+ is reference category)				
<$10,000	2.07[†] (1.30, 3.30)	1.99[†] (1.07, 3.70)	1.72[†] (1.01, 2.94)	2.16[†] (1.10, 4.22)
$10,000 to $19,999	1.41 (0.87, 2.29)	1.58 (0.83, 3.01)	1.53 (0.88, 2.66)	1.80 (0.90, 3.62)
$20,000 to $39,999	1.35 (0.86, 2.13)	1.16 (0.62, 2.20)	1.78[†] (1.06, 2.98)	1.16 (0.58, 2.31)

Each CAGE question served as a dependent variable in the analyses. Odds ratios are adjusted for alcohol use disorders and consumption patterns. Odds ratios represent the increased likelihood that a patient with the factor (e.g., a male) responded "Yes" to the CAGE question, compared with a patient without the factor (e.g., a female).
[†]*Odds ratio significant at* p < .05.

cohol use problems for men and women. Proportionately, more women are lifetime abstainers, while more men have experienced problems related to their drinking.[5] More men are subthreshold for an alcohol use disorder, and we would expect that a large proportion of men without a disorder would endorse each CAGE question.

Mexican–American patients were less likely to endorse each CAGE question "yes" than white or African–American patients after adjustment for the same covariates. This finding is consistent with research on patterns of alcohol consumption among Hispanics, in whom alcohol use is associated with celebration and high-volume drinking over limited periods of time.[13] This type of drinking pattern may be normative. Thus, for others to have complained about the patient's alcohol use or for the patient to report having felt guilty or having the need to cut down on drinking, or having consumed alcohol in the morning after an evening of heavy alcohol consumption, is less likely.

The adjusted models predicting an alcohol use disorder using the CAGE questions showed many inconsistencies. The *eye-opener* question was not retained in the adjusted models for men and white patients. One explanation for this finding is that drinking in the morning after an evening of heavy consumption is not uncommon among young adult men.[14] Thus, this question may be poorly spe-

cific as many men who do not meet criteria for an alcohol use disorder have engaged in this pattern of alcohol use.

The *annoyed* question appears to be poorly sensitive for Mexican Americans. Again, this finding is consistent with patterns of alcohol consumption among Hispanics.[13] The drinking pattern may be normative, and others are less likely to complain about alcohol use. The *annoyed* question was not retained in the models for white or African–American patients and adds little information for predicting an alcohol use disorder because of its correlation with the other questions. The *guilty* question was not retained in the any of the adjusted models, adding little to the model when these other questions are included.

The brevity of the CAGE and its ease of interpretation suggest that it should not be abandoned entirely. In clinical practice, the role of the CAGE in screening might be greatest for patients who respond positively to *any* of the CAGE questions. This argues for a lower cutpoint (a single positive response) to minimize the false-negative rate. The CAGE, though, may do little to rule out an alcohol use problem, in particular for Mexican–American patients in whom the sensitivity is low. In contrast, the *cut-down* question was consistently predictive of an alcohol use disorder. The use of the *cut-down* question in combination with queries on quantity or frequency of alcohol use may offer an acceptable approach to screening.

Table 5. Adjusted Odds Ratios (95% Confidence Limits) from Logistic Regression Models, with an Alcohol Use Disorder Predicted by CAGE Questions*

CAGE Question	All Subjects	Men	Women	Whites	African Americans	Mexican Americans
Cut-down	7.76 (5.42, 11.12)	5.92 (3.49, 10.66)	8.29 (5.04, 13.64)	10.15 (5.86, 17.57)	11.56 (6.62, 20.17)	5.02 (2.56, 9.84)
Annoyed	—	—	2.24 (1.11, 4.51)	—	—	—
Guilty	—	—	—	—	—	—
Eye-opener	2.19 (1.34, 3.59)	—	2.29 (1.06, 4.92)	—	4.28 (1.88, 9.75)	4.71 (1.47, 15.11)

Models were estimated using forward entry criteria for selection of CAGE questions. Questions that did not enter are indicated by dashes. Unless serving as a subgroup selection variable, odds ratios are adjusted for annual income, gender, and race or ethnicity of patient.

This study is limited by use of a single clinical site and cross-sectional design. The prognostic validity of the CAGE is unknown. Furthermore, at least one other study has suggested that the sensitivity of the CAGE is higher when it is administered before the questions about quantity or frequency of alcohol consumption (the CAGE questions followed consumption questions in the present study).[15] We do not know whether the effects of such administration would differentially impact sensitivity across patient subgroups.

Recently, recommendations have been made to systematically incorporate the CAGE questions into primary care practices, because physicians use the screen inconsistently.[16] The CAGE offers many advantages, including brevity and ease of recall. This study supports the need for brief and unbiased screening tests for alcohol use disorders in primary care patients. Future research might consider the following criteria in developing and evaluating screening tests: the screen must be brief (brevity); it should minimize false-negative results while not resulting in a high rate of false-positive results (accuracy); it should perform similarly for men and women of different racial or ethnic backgrounds, ages, and socioeconomic status (absence of subgroup bias); it should show consistency across varied clinical settings, where the prevalence and spectrum of alcohol use problems may vary (absence of spectrum bias); and results from the screen should inform treatment (clinical utility). The CAGE appears to fulfill some, but not all of these criteria.

REFERENCES

1. Magruder HK, Durand AM, Frey KA. Alcohol abuse and alcoholism in primary health care settings. J Fam Pract. 1991;32:406–13.

2. Allen JP, Eckardt MJ, Wallen J. Screening for alcoholism: techniques and issues. Public Health Rep. 1988;103:586–92.

3. Allen JP, Maisto SA, Connors GJ. Self-report screening tests for alcohol problems in primary care. Arch Intern Med. 1995;155:1726–30.

4. Beresford TP, Blow FC, Brower KJ, Singer K. Clinical applications: screening for alcoholism. Prev Med. 1988;17:653–63.

5. Volk RJ, Steinbauer JR, Cantor SB, Holzer CE. Screening for "at risk" drinking in primary care patients of different racial/ethnic backgrounds. Addiction. 1997;92:197–206.

6. Ewing JA, Rouse BA. Identifying the hidden alcoholic. In: The 29th International Congress on Alcohol and Drug Dependence. Sydney, Australia; 1970.

7. Ewing JA. Detecting alcoholism: the CAGE Questionnaire. JAMA. 1984;252:1905–7.

8. Grant BF, Hasin D. The Alcohol Use Disorder and Associated Disabilities Interview Schedule (AUDADIS). Rockville, Md: National Institute on Alcohol Abuse and Alcoholism; 1992.

9. Grant BF, Harford TC, Dawson DA, Chou PS, Pickering RP. The Alcohol Use Disorder and Associated Disabilities Interview Schedule (AUDADIS): reliability of alcohol and drug modules in a general population sample. Drug Alchol Depend. 1995;39:37–44.

10. Cahalan DR, Roizen R, Room R. Alcohol problems and their prevention: public attitudes in California. In: Room R, Sheffield S, eds. The Prevention of Alcohol Problems: Report of a Conference. Sacramento, Calif: Health and Welfare Agency, Office of Alcoholism; 1974:354–403.

11. Cherpitel CJS. Drinking patterns and problems associated with injury status in emergency room admissions. Alcohol Clin Exp Res. 1988;12:105–10.

12. Bush B, Shaw S, Cleary P, Delbanco TL, Aronson MD. Screening for alcohol abuse using the CAGE Questionnaire. Am J Med. 1987;82:231–87.

13. Caetano R. Acculturation, drinking and social settings among U.S. Hispanics. Drug Alcohol Depend. 1987;19:215–26.

14. Johnston LD, O'Malley PM, Bachman JG. National Survey Results on Drug Use From the Monitoring the Future Study, 1975–1992, Vol 2; College Students and Young Adults. Rockville, Md: National Institute on Drug Abuse; 1993. NIH publication 93–3598.

15. Steinweg DL, Worth H. Alcoholism: the keys to the CAGE. Am J Med. 1993;94:520–3.

16. Wenrich MD, Paauw DS, Carline JD, Curtis JR, Ramsy PG. Do primary care physicians screen patients about alcohol intake using the CAGE questions? J Gen Intern Med. 1995;10:631–4.