Alcohol Abuse and Dependence in Latinos Living in the United States

Validation of the CAGE (4M) Questions

Richard Saitz, MD, MPH; Mark F. Lepore, BA; Lisa M. Sullivan, PhD; Hortensia Amaro, PhD; Jeffrey H. Samet, MD, MA, MPH

Background: Brief alcoholism screening questionnaires have not been adequately studied in the rapidly growing Latino population living in the United States.

Objective: To assess (1) the prevalence of alcoholism and (2) the performance of 2 alcohol screening instruments in Latinos.

Subjects and Methods: We performed a cross-sectional interview study in an urban teaching hospital-based primary care practice. Consecutive self-identified Latino subjects provided informed consent. All subjects were interviewed in English or Spanish using 2 alcoholism screening tools, the CAGE (or the Spanish version, the 4M), and the Alcohol Use Disorders Identification Test, and a criterion standard for the diagnosis of alcohol abuse and dependence, the Composite International Diagnostic Interview.

Results: Of 210 subjects interviewed, 36% had a lifetime diagnosis of alcohol abuse or dependence by the criterion standard. Thirty-one percent were currently drinking hazardous amounts of alcohol. A CAGE (4M) score of 1 or more was 92% sensitive and 74% specific, and a score of 2 or more was 80% sensitive and 93% specific for a lifetime diagnosis of alcohol abuse or dependence. CAGE (4M) scores of 0, 2, 3, and 4 were associated with likelihood ratios (0.1, 4.8, 18.5, and 36.8, respectively) that resulted in substantial changes from pretest (36%) to posttest probability (to 6%, 73%, 91%, and 95%, respectively) of a diagnosis of alcohol abuse or dependence. At the standard cutoff point, the Alcohol Use Disorders Identification Test detected only 51% of subjects with alcohol disorders.

Conclusions: In Latinos in primary care settings, alcohol abuse and dependence are common and the CAGE (4M) is a brief, valid, screening tool for detecting alcohol use disorders.

Arch Intern Med. 1999;159:718-724

In the United States, alcoholism is a leading cause of death and costs $148 billion each year.1,2 Rapid, accurate screening instruments can detect alcohol problems in primary care settings.3-6 Brief interventions for these problems positively impact alcohol consumption, morbidity, and mortality.7-11

Studies have not delineated the prevalence of alcoholism in Latinos in primary care settings in the United States. In population-based surveys, however, heavy drinking is as common in Latinos as in African Americans and non-Latino whites.3 But serious consequences of heavy drinking are more common in Latinos than in other ethnic groups.12

Few alcoholism screening tests have been evaluated for use in Latinos or Spanish speakers and none have been tested in the largest groups of Latinos living in the United States. The CAGE has been validated in Spain.13-15 (CAGE acronym arises from key concepts contained in each of the 4 questionnaire items: Have you ever felt you should cut down on your drinking? Have people annoyed you by criticizing your drinking? Have you felt bad or guilty about your drinking? Have you ever had a morning eye-opener (used alcohol first thing in the morning to steady your nerves or get rid of a hangover?) Other more lengthy instruments have been studied in Mexico, and in Mexican Americans.16-18 But screening tests developed and tested outside the United States may not be valid in Latinos living in the United States.15,19 These screening tests rely in part on patients’ perceptions of their drinking, which may differ according to sex, ethnic origin, and acculturation.20

Therefore, we tested 2 hypotheses: (1) that the prevalence of alcohol abuse or dependence in Latinos visiting a primary care center would be high, and (2) that screening tests developed and validated in non-Latinos would not be valid in a diverse La-
SUBJECTS AND METHODS

SUBJECTS

Eligible subjects considered themselves to be Latino. Patients visiting an urban teaching hospital-based primary care center were approached after registration for a medical visit, while waiting for their physician. The study was approved by the Human Studies Committee of the Boston Medical Center, Boston, Mass, and all subjects provided informed consent.

DATA COLLECTION

Data were collected by interview with 1 of 3 bilingual staff researchers, 2 of whom were Latino. After being asked questions regarding demographics, ethnic origin, and the short acculturation scale, the alcohol section of the interview began with the 4 CAGE questions (scored 0-4). The Spanish CAGE questions (the 4M) were derived from those validated in a primary care setting in Spain, and were modified based on the focus group comments of Dominicans and Puerto Ricans living in the clinic’s catchment area. The 4M (Spanish version of CAGE) questions were:

- Ha tenido Usted alguna vez la impresión de que deberías beber menos?
- Le ha molestado alguna vez la gente criticándole su forma de beber?
- Alguna vez lo primero que ha hecho por la mañana ha sido beber para calmar los nervios o para librarase de una goma (una resaca)?
- Ha tenido Usted alguna vez la impresión de que debería beber menos?

The second screening tool was the Alcohol Use Disorders Identification Test (AUDIT) (10 items scored 0-40). The Spanish version was slightly modified from a published version to improve comprehension based on focus group comments.

Finally, all subjects completed the Alcohol Module of the Composite International Diagnostic Interview Version 2.0, a criterion standard that yields a Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) diagnosis of alcohol abuse or dependence.

DEFINITIONS

Subjects ever having had alcohol abuse or dependence (a lifetime diagnosis), reported the symptoms required for diagnosis during a 12-month period anytime during their lives; subjects with current diagnoses reported the required symptoms within the past year.

Hazardous drinking amounts (>14 standard drinks per week [7 for women] or >4 per occasion [3 for women]) were assessed using the first 3 questions of the AUDIT.

STATISTICAL ANALYSIS

Analyses were performed using PC SAS statistical software (Version 6.12). Sociodemographic characteristics were compared among those with alcohol diagnoses or not using t tests and χ² tests. Estimates and 95% confidence intervals of the sensitivity, specificity, and likelihood ratios were calculated using published formulas. Receiver operating characteristic (ROC) curves were constructed and the areas under ROC curves were estimated along with SEs and 95% confidence intervals. To evaluate whether the ROC curves differed by subject or interviewer characteristics, we (1) developed separate ROC curves on subgroups of subjects stratified according to the characteristic of interest, (2) visually inspected the separate ROC curves overlaid, and (3) tested for a significant difference between the areas under the ROC curves.

PREVALENCE OF ALCOHOL DIAGNOSES

Based on the diagnostic criterion standard, the Composite International Diagnostic Interview, 76 (36%) of 210 subjects met DSM-IV criteria for ever having had alcohol abuse or dependence (a lifetime diagnosis). Lifetime alcohol abuse or dependence was more common in men than women (53% vs 17%; P = .001), in Puerto Ricans and Central Americans than in Dominicans (47%, 41%, and 22%, respectively; P = .01), and in the small minority born on the US mainland (77% vs 23%; P = .008). Subjects with alcohol abuse or dependence had lived in the United States longer (mean, 18 vs 15 years; P = .04). Age, acculturation, and education were similar in individuals with and without a lifetime alcohol diagnosis.

Sixteen (8%) of 210 subjects met DSM-IV criteria for current alcohol abuse or dependence. Sixty-five (31%) of the 210 subjects were currently drinking hazardous amounts.

The prevalence of a lifetime diagnosis of alcohol abuse or dependence, demographics, and acculturation were similar regardless of interviewer (data not shown).
SCREENING FOR A LIFETIME ALCOHOL DISORDER

The operating characteristics of the CAGE (4M) screening tool compared with the DSM-IV diagnosis of lifetime alcohol abuse or dependence appear in Table 2. CAGE (4M) scores of 1 or greater (achieved in 105 [50%] of 210 subjects) and 2 or greater (achieved in 71 [34%] of 210 subjects) were reasonably sensitive (92% and 80%, respectively) and specific (74% and 93%, respectively).

Likelihood ratios associated with CAGE (4M) scores appear in Table 2. CAGE (4M) scores of 0, 2, 3, and 4 were associated with influential likelihood ratios. Only 16% of subjects had a CAGE (4M) score equal to 1, which was associated with little change from pretest to posttest probability. A CAGE (4M) score of 0 was associated with a likelihood ratio of 0.1 and a posttest probability of 6% (given the observed prevalence of 36%). CAGE (4M) scores of 2 or more were associated with likelihood ratios of 4.8 or greater and posttest probabilities of 73% to 95%. The ROC curve in Figure 1 shows the tradeoffs in sensitivity and specificity for each possible cutoff point of the CAGE (4M). A cutoff score of 2 minimizes the sum of false positives and false negatives.

We also examined the sensitivity and specificity of the 4 individual CAGE (4M) items (Table 3). The “cutting down” (“menos”) question was the most sensitive item. The “eye-opener” (“manana”) item, answered in the affirmative by only 16% of subjects, was the most specific but least sensitive. The positive likelihood ratio was 20.5, and the negative likelihood ratio, 0.1 for the eye-opener item.

The sensitivity and specificity of the AUDIT for a lifetime diagnosis of alcohol abuse or dependence appear in Table 4. The AUDIT scores of 1 or greater were 89% sensitive but only 50% specific. Although the specificity rises with an increase in the AUDIT score, the sensitivity drops to only 51% at a score of 8 or greater, the standard clinical cutoff score.16 Scores of 8 or more, achieved by 45 (21%) of 210 subjects, were associated...
with influential likelihood ratios, but at these cutoff scores almost half the cases would remain undetected. The ROC curve in Figure 2 shows the tradeoffs in sensitivity and specificity for each possible cutoff point of the AUDIT.

SCREENING FOR A CURRENT ALCOHOL DISORDER

A CAGE (4M) score of 1 or greater was 100% sensitive for current alcohol abuse or dependence (Table 5). Likelihood ratios associated with a current disorder were less influential than for a lifetime diagnosis. Although only 54% specific, the posttest probability nearly doubled at a score of 1 from 6% to 15%. A cutoff score of 2 provided minimal further gains in posttest probabilities.

Likelihood ratios associated with the AUDIT for a current disorder were also less influential than for a lifetime diagnosis of alcohol abuse or dependence. The AUDIT scores of less than 8 were associated with likelihood ratios that resulted in a decrease in the probability of current alcoholism (from the 8% prevalence observed in the sample to <4%) (Table 6). Scores of 12 or greater (achieved by 13% of the sample) were associated with a moderately large likelihood ratio (6.7) and a large increase from pretest (8%) to posttest (36%) probability.

**Table 4. Sensitivity, Specificity, Likelihood Ratio, and Posttest Probability of AUDIT Scores and Lifetime Alcohol Abuse or Dependence**

<table>
<thead>
<tr>
<th>AUDIT Score</th>
<th>Sensitivity, %</th>
<th>Specificity, %</th>
<th>Likelihood Ratio</th>
<th>Posttest Probability, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>...</td>
<td>...</td>
<td>0.2 (0.1-0.4)</td>
<td>11 (4-18)</td>
</tr>
<tr>
<td>1-3</td>
<td>89 (83-96)</td>
<td>50 (42-58)</td>
<td>0.6 (0.5-0.7)</td>
<td>24 (13-34)</td>
</tr>
<tr>
<td>4-7</td>
<td>70 (59-80)</td>
<td>86 (80-92)</td>
<td>1.9 (0.9-3.8)</td>
<td>52 (33-71)</td>
</tr>
<tr>
<td>8-11</td>
<td>51 (40-63)</td>
<td>96 (92-99)</td>
<td>5.7 (1.9-16.9)</td>
<td>77 (56-97)</td>
</tr>
<tr>
<td>≥12</td>
<td>34 (24-45)</td>
<td>99 (96-100)</td>
<td>23.0 (5.6-94.2)</td>
<td>93 (83-100)</td>
</tr>
</tbody>
</table>

*Lifetime alcohol abuse or dependence was determined using the criterion standard Composite International Diagnostic Interview (see the “Subjects and Methods” section of the text). Sensitivity and specificity refer to characteristics of an Alcohol Use Disorders Identification Test (AUDIT) score greater than or equal to the lowest score in the range shown. Likelihood ratios and posttest probabilities correspond to AUDIT scores in the score range shown. Values within parentheses are 95% confidence intervals. Ellipses indicate not applicable.†Posttest probabilities are calculated based on a pretest probability of 36% observed in the sample.

**Table 5. Sensitivity, Specificity, Likelihood Ratio, and Posttest Probability of CAGE (4M) Scores and Current Alcohol Abuse or Dependence**

<table>
<thead>
<tr>
<th>CAGE (4M) Score</th>
<th>Sensitivity, %</th>
<th>Specificity, %</th>
<th>Likelihood Ratio</th>
<th>Posttest Probability, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>...</td>
<td>...</td>
<td>0</td>
<td>0‡</td>
</tr>
<tr>
<td>1</td>
<td>100</td>
<td>54 (47-61)</td>
<td>0.8 (0.2-2.9)</td>
<td>6 (0-14)</td>
</tr>
<tr>
<td>2</td>
<td>88 (71-100)</td>
<td>71 (64-77)</td>
<td>2.2 (0.9-5.6)</td>
<td>15 (2-29)</td>
</tr>
<tr>
<td>3</td>
<td>63 (39-86)</td>
<td>82 (77-87)</td>
<td>2.6 (1.0-6.6)</td>
<td>17 (2-33)</td>
</tr>
<tr>
<td>4</td>
<td>38 (14-61)</td>
<td>92 (88-96)</td>
<td>4.5 (2.1-10.0)</td>
<td>27 (9-46)</td>
</tr>
</tbody>
</table>

*4M represents the Spanish version of CAGE, which is a questionnaire for alcoholism screening. Lifetime alcohol abuse or dependence was determined using the criterion standard Composite International Diagnostic Interview (see the “Subjects and Methods” section of the text). Sensitivity and specificity refer to characteristics of a CAGE (4M) score greater than or equal to the score shown. Likelihood ratios and posttest probabilities correspond to CAGE (4M) scores equal to the score shown. Values within parentheses are 95% confidence intervals. Ellipses indicate not applicable.†Posttest probabilities are calculated based on a pretest probability of 8% observed in the sample.‡One hundred five of 210 subjects had a cage (4M) score of 0. None of them met diagnostic criteria for current alcohol abuse or dependence.

**Table 6. Sensitivity, Specificity, Likelihood Ratio, and Posttest Probability of AUDIT Scores and Current Alcohol Abuse or Dependence**

<table>
<thead>
<tr>
<th>AUDIT Score</th>
<th>Sensitivity, %</th>
<th>Specificity, %</th>
<th>Likelihood Ratio</th>
<th>Posttest Probability, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>...</td>
<td>...</td>
<td>0</td>
<td>0‡</td>
</tr>
<tr>
<td>1-3</td>
<td>100</td>
<td>39 (32-46)</td>
<td>0.2 (0.1-1.3)</td>
<td>2 (0-5)</td>
</tr>
<tr>
<td>4-7</td>
<td>94 (82-100)</td>
<td>71 (64-77)</td>
<td>0.5 (0.1-3.2)</td>
<td>4 (0-11)</td>
</tr>
<tr>
<td>8-11</td>
<td>88 (71-100)</td>
<td>84 (79-89)</td>
<td>3.7 (1.4-10.1)</td>
<td>24 (3-44)</td>
</tr>
<tr>
<td>≥12</td>
<td>63 (39-86)</td>
<td>91 (87-95)</td>
<td>6.7 (3.8-12.0)</td>
<td>36 (18-53)</td>
</tr>
</tbody>
</table>

*Lifetime alcohol abuse or dependence was determined using the criterion standard Composite International Diagnostic Interview (see the “Subjects and Methods” section of the text). Sensitivity and specificity refer to characteristics of an Alcohol Use Disorders Identification Test (AUDIT) score greater than or equal to the lowest score in the range shown. Likelihood ratios and posttest probabilities correspond to AUDIT scores in the score range shown. Values within parentheses are 95% confidence intervals. Ellipses indicate not applicable.†Posttest probabilities are calculated based on a pretest probability of 8% observed in the sample.‡Seventy-five of 210 subjects had an AUDIT score of 0. None of them met diagnostic criteria for current alcohol abuse or dependence.

**Table 2. Receiver operating characteristic curve for Alcohol Use Disorders Identification Test (AUDIT) scores and the identification of lifetime alcohol abuse and dependence. Selected AUDIT scores appear adjacent to the curve.**

©1999 American Medical Association. All rights reserved.
Figure 3. Receiver operating characteristic curve for CAGE, an alcoholism screening questionnaire (Spanish version [4M]) scores and the identification of lifetime alcohol abuse or dependence stratified by sex. Scores appear adjacent to the curve and in men are represented by open squares; in women by solid circles.

Although some researchers have reported lack of sensitivity, our data revealed that the CAGE (4M) was sensitive for early identification of hazardous and problematic drinkers. As in prior studies, the AUDIT was sensitive for current disorders but often missed past alcohol problems, which are important to identify in the primary care setting. Because of the AUDIT's lack of sensitivity and its length, it is less desirable as a physician-administered screening tool in primary care settings. To further augment the sensitivity of the CAGE (4M), a few questions regarding the quantity and frequency of usual alcohol intake should be asked after (not before) asking the CAGE (4M) questions.

Although the CAGE had been validated in many populations, we set out to validate the CAGE in Latinos because we suspected that a screening test that relied on a patient's perception of their drinking might not perform well. Acculturation, country of origin, and sex influence social norms and drinking patterns, and, therefore, perceptions of harmful drinking. For example, the concept of machismo, or manliness, which includes the ability to drink large amounts of alcohol frequently without showing intoxication, may make some Latino men less likely to recognize a problem.

We found the CAGE (4M) to be valid and did not detect any difference in test performance in subjects of different origins, educational levels, acculturation, or sex.

However, our results should not be interpreted to imply that the CAGE will be valid in all cultures. Nelson et al. reported that although 39% of Vietnamese immigrants reported alcohol use, none answered any of the CAGE questions in the affirmative. Testing of the CAGE against a criterion standard remains important when considering its use in new populations.

In deciding an appropriate cutoff score for the CAGE (4M) (ie, a score of 1 or 2), both the frequency and the consequences of false positives and false negatives should be considered. A score of 1 or greater was the most sensitive; a score of 2 or higher greatly increased the posttest probability at the expense of a decline in sensitivity. Given the greater consequences of missing the diagnosis, and the ease with which a false positive could be clarified, we agree with prior recommendations that the cutoff score of 1 or greater be used for screening.

Several limitations should be considered in interpreting and applying the results presented. The interviews were done by trained staff researchers. These methods may have yielded results different from those one might see in clinical practice. However, the CAGE and AUDIT have been administered in a variety of health care settings and formats (written, interview, and computer), and by different interviewers, with similar results in many other populations.

Generalizability may be limited to populations similar to those we studied: minimally acculturated urban dwellers in the northeastern United States visiting a primary care center, and of Caribbean, Central, and South...
Accepted for publication July 15, 1998.

Mr Lepore and Drs Saitz, Amaro, and Samet were supported in this work by grant T15 SP07773-01 from the Center for Substance Abuse Prevention Faculty Development Program, Substance Abuse Mental Health Services Administration, US Department of Health and Human Services, Washington, DC. Dr Saitz is a Robert Wood Johnson Foundation Generalist Physician Faculty Scholar.

Our results revealed the high prevalence of alcohol abuse and dependence in Latino subjects and validate the CAGE (4M) questionnaire in Latinos. We also confirmed that the AUDIT is insensitive for past alcohol problems. Our results demonstrate that current recommendations to screen for alcohol abuse in primary care settings are applicable to Latinos living in the United States and that the CAGE (4M) questions can be effectively used to achieve this goal.

REFERENCES


