ORIGINAL ARTICLE

Alcohol Screening in Young Persons Attending a Sexually Transmitted Disease Clinic

Comparison of AUDIT, CRAFFT, and CAGE Instruments

Robert L. Cook, MD, MPH,[1] Tammy Chung, PhD,[2] Thomas M. Kelly, PhD,[2] Duncan B. Clark, MD, PhD[2]

[1]Division of General Internal Medicine, School of Medicine, University of Pittsburgh, Pittsburgh, PA, USA; [2]Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA, USA.

OBJECTIVE: To compare the ability of 3 brief alcohol screens (Alcohol Use Disorders Identification Test [AUDIT], CRAFFT, and CAGE) to identify adolescents and young adults with a current alcohol use disorder (AUD) and to determine whether there are gender-based or race-based differences in screening performance.

DESIGN, PARTICIPANTS, AND SETTING: Cross-sectional study of 358 young persons (55% males; 49% blacks; age range, 15–24 years; mean age, 20.6 years) who were attending an urban clinic for sexually transmitted diseases and reported alcohol use during the past year.

MEASUREMENTS: Receiver operating characteristic (ROC) curve analysis was used to determine the ability of the 3 screens to discriminate between participants with and without AUDs detected in the Structured Clinical Interview for DSM-IV (SCID).

RESULTS: One third (33%) of participants met Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV) criteria for a current AUD (24% with alcohol abuse and 9% with alcohol dependence). The AUDIT performed best at a cut score of 9 (sensitivity, 0.76; specificity, 0.79), CRAFFT at a cut score of 2 (sensitivity, 0.94; specificity, 0.33), and CAGE at a cut score of 1 (sensitivity, 0.69; specificity, 0.63). The AUDIT had the best overall performance (area under the curve [AUC], 0.84), followed closely by CRAFFT (AUC, 0.79) and then CAGE (AUC, 0.70). Performance of screens did not differ by gender. The AUDIT performed slightly better in whites than blacks, but no race-based differences were observed for the CAGE or CRAFFT.

CONCLUSIONS: Clinicians should use the AUDIT or CRAFFT, rather than the CAGE, to screen young persons for AUDs. The AUDIT performs best, but its length may limit its utility in this setting. The CRAFFT is a suitable alternative, with excellent sensitivity and no gender-based or race-based differences.

KEY WORDS: alcoholism; alcohol drinking; screening; adolescents.
DOI: 10.1111/j.1525-1497.2005.40052.x
J GEN INTERN MED 2005; 20:1–6.

Alcohol consumption is a major contributor to risky behaviors and adverse health outcomes in adolescents and young adults. Motor vehicle crashes, homicides, suicides, and unintentional injuries are the 4 leading causes of death in individuals aged 15–20 years, and alcohol plays a substantial role in many of these events.[1] In addition, alcohol consumption is associated with risky sexual behavior, sexually transmitted diseases (STDs), and school and social problems.[1]

The U.S. Preventive Services Task Force has encouraged clinicians to screen young persons for problematic drinking.[2] Screening may be offered in the primary care setting when individuals undergo a routine physical examination or seek treatment for acute problems, or it may be offered in specialized medical settings, such as emergency departments or STD clinics, where the prevalence of alcohol problems is reported to be particularly high.[3,4] Although universal screening for problem drinking has been recommended,[2] medical staff do not consistently comply with this recommendation.[5,6] One barrier to universal screening is the relative absence of evidence regarding the most accurate and efficient brief alcohol screening method to use in a particular medical setting and with a specific population.

As defined in the fourth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV), alcohol use disorders (AUDs) include alcohol abuse and alcohol dependence. Abuse is characterized by substance-related impairment in fulfilling major role obligations, hazardous use, legal problems related to use, and interpersonal problems related to use. Dependence is characterized by symptoms such as tolerance, withdrawal, and other indicators of a compulsive pattern of use. Diagnosis by strict DSM-IV criteria requires a detailed clinical interview that is often complicated and impractical to carry out in primary care or community-based settings. Therefore, it is important to identify brief screening tools that can accurately identify persons who may have significant drinking problems.

Several brief screening measures are available for use in medical settings. The most popular alcohol screening measure in primary care is a 4-item survey called CAGE.[7] A second measure, the Alcohol Use Disorders Identification Test (AUDIT), has been studied in a variety of settings but may be less familiar to many clinicians.[8] A third screening tool, called CRAFFT, was developed specifically to screen for problematic substance use in teens.[9] The AUDIT and CAGE have been compared in several studies of adolescents and young adults, and results indicate that the AUDIT has a better overall performance than the CAGE in discriminating persons with and without an AUD.[10–12] Although the CRAFFT has been assessed in fewer studies, results among adolescents have been promising. In a recent study, Knight et al. (2003) evaluated the performance of several screening tests in a sample of 538 teenagers aged 14–18 recruited from primary care clinics.[12] In that study, the CRAFFT had the best overall screening performance, followed by the AUDIT and then the CAGE. The CRAFFT has not been tested among a broader age spectrum (e.g., 15–24 years), and

Accepted for publication June 1, 2004

None of the authors has any conflict of interest relevant to this manuscript or the data.

This research was presented in part at the June 22–26, 2003 annual meeting of the Research Society on Alcoholism, Ft. Lauderdale, FL.

Address correspondence and requests for reprints to Dr. Cook: 230 McKee Place, Suite 600, Pittsburgh, PA 15213 (e-mail: cookrl@upmc.edu).

See editorial by O'Connor, p. 96–97.

its characteristics might vary in a clinical setting where the prevalence of alcohol problems is greater.

With any screening measure, it is important to have evidence that the screening tool functions with similar accuracy, regardless of gender, race, age, or setting. Some research suggests that alcohol screens such as the CAGE or the AUDIT have a lower sensitivity in females,[13–15] although a review of 9 studies involving samples of adults concluded that the overall performance of the CAGE and AUDIT generally did not differ by gender.[16] We are not aware of previous studies that compared the performance of the CRAFFT by race or gender. Although the performance of the CAGE, CRAFFT, and AUDIT have been examined in either adult or adolescent medical patients, the best brief alcohol screen for use in a clinic that serves both adolescents and young adults remains to be identified. It is also unclear whether the performance of alcohol screens would differ in a clinic serving a higher-risk population (such as an STD clinic) compared to a primary care clinic, where the prevalence of alcohol problems is likely to be much lower.

Because little is known about the relative performance of the AUDIT, CRAFFT, and CAGE measures across gender and race in adolescents and young adults, our primary objective was to explore this issue by comparing the ability of these 3 brief alcohol screens to identify the presence of a current (12-month) DSM-IV-defined AUD in young men and women attending a public STD clinic.

METHODS

Participants and Procedures

Young men and women were eligible for the study if they met the following criteria: they were attending an urban STD clinic in Pittsburgh, PA for a new clinical assessment; they were 15–24 years old; and they received a full clinical assessment, along with routine diagnostic testing at the clinic. Pregnant women were excluded, because pregnancy could affect substance use behavior. The majority of participants were recruited directly by research assistants when they attended the clinic for a screening test or new problem during 1999–2001. During this time, 763 eligible persons were approached and 328 agreed to participate (43%). We did not specifically ask nonparticipants about reasons for nonparticipation, but the majority of those who did not participate indicated a lack of interest or time. Participants were compared to nonparticipants on gender, race, alcohol consumption, and clinical impression of an STD. The only significant difference between the two groups was that females were more likely to participate than males (51% vs 37%; $P<.001$). An additional 71 persons were recruited into the study after self-referral in response to a brochure distributed at the clinic. These subjects met the same eligibility criteria as those recruited by the research assistants. Participants in the 2 recruitment groups did not differ in terms of demographic characteristics, the rates of diagnosis of an alcohol or cannabis use disorder, or the rates of diagnosis of an STD. Each participant received and signed an informed consent form approved by the University of Pittsburgh Institutional Review Board.

Participants first were asked to complete a written questionnaire that included questions from the AUDIT, CRAFFT, and CAGE measures (described below). Written instruments for the alcohol screening measures were used so that the research assistants would not have the results of the screens prior to the follow-up interview, and because the results of alcohol screening instruments typically do not vary according to data collection method.[17] Approximately 10% of participants did complete the screening tests by interview, including some who were noted to have literacy problems during the consent process, although scoring was not completed until after all data were collected.

Next, participants completed a structured interview that either took place in person on the day of enrollment or was scheduled to take place in person or via the telephone within 30 days of enrollment. The interviews included items to assess demographic characteristics and alcohol and drug use behavior, and were conducted by 2 female research assistants, each of whom had completed extensive training in the structured clinical assessment of alcohol and drug use. After completing the interview, each participant received a small cash incentive of $20–$40, together with information about alcohol and drug treatment options available in the community.

Measures

The AUDIT consists of 10 items that cover the "past year" time frame.[8] Response options for each item range from 0 to 4, resulting in a total possible score of 40. Adults who have a score of 8 or more are considered to have an AUD.[8] In contrast to most other brief alcohol screens, which assume that the level of alcohol consumption has already been queried, the AUDIT includes 3 items that assess the quantity and frequency of alcohol use. Some studies of the use of the AUDIT in adolescents and adults suggest that the AUDIT consists of 2 subscales, one representing "level of consumption" and the other representing "alcohol-related problems."[18–20] In our study sample of 358 participants, a preliminary analysis indicated that the AUDIT total score showed better overall performance than did the AUDIT level of consumption subscale (χ^2 [$df=1$, $n=358$]$=9.31$; $P<.01$) or the AUDIT alcohol-related problems subscale (χ^2 [$df=1$, $n=358$]$=9.72$; $P<.01$). Therefore, our study results focus on the AUDIT total score.

CRAFFT is a mnemonic that cues 6 items covering a "past year" time frame.[9] We modified the items to query alcohol only, rather than "alcohol or drug use" as in the original measure. The 6 CRAFFT items are: have you ridden in a *C*ar driven by someone (including yourself) who had been drinking? Do you use alcohol to *R*elax, feel better about yourself, or fit in? Do you use alcohol while you are by yourself, *A*lone? Do you *F*orget things you did while using alcohol? Do your family or *F*riends tell you that you should cut down on your drinking? Have you gotten into *T*rouble while using alcohol? The CRAFFT has been found to perform best at a cut score of 2 when used to identify adolescents with a DSM-IV substance use disorder in a medical clinic setting.[12]

CAGE is a mnemonic that cues 4 items covering a "past year" time frame[7]: have you felt that you should *C*ut down on your drinking? Have people *A*nnoyed you by criticizing your drinking? Have you felt bad or *G*uilty about your drinking? Have you had a drink first thing in the morning (*E*ye-opener) to steady your nerves or to get rid of a hangover? Positive responses to 1 or 2 items result in a positive screen.[7,10,12]

To determine whether a DSM-IV-defined AUD was present in a participant during the past 12 months, we used a modified version of the Structured Clinical Interview for DSM-IV Sub-

stance Use Disorders (SCID).[21] This version of the SCID had been modified to accommodate developmental considerations in the assessment of substance-related symptoms in adolescents and young adults (e.g., missing school because of alcohol use) and had been found to have moderate to high interrater reliability and good concurrent validity.[22,23]

Data Analyses

In addition to determining the prevalence of past year DSM-IV-defined AUDs in our total sample, we determined the prevalence according to gender and race. For analyses by gender, we included all 358 participants. For analyses by race, we compared results in whites (165 participants) with those in blacks (177 participants) and excluded results in participants from "other" ethnic backgrounds (16 participants). We did not stratify by age, because preliminary analyses indicated that age was not correlated with scores on any of the screens (age with AUDIT: $r=-.023$, $P=.630$; age with CRAFFT: $r=.054$, $P=.31$; and age with CAGE: $r=-.001$, $P=.989$).

Sensitivity (probability of a positive screening result in those with a diagnosis) and specificity (probability of a negative screening result in those without a diagnosis) were computed to determine the performance of the 3 screens against the presence of a DSM-IV-defined AUD as determined by the SCID (i.e., no diagnosis vs presence of either alcohol abuse or alcohol dependence). A receiver operating characteristic (ROC) curve, which describes changes in a screen's sensitivity (true-positive rate) as a function of the false-positive rate (1 − specificity), was plotted for each possible cut score for each screen using the AccuROC version 2.5 software program (Accumetric Corporation, Montreal, Quebec). For each screening test and possible cut point, we also calculated the positive predictive value (PPV; the probability of a diagnosis in those with a positive screening test) and the negative predictive value (NPV; the probably of the lack of a diagnosis in those with a negative screening result).

AccuROC was used to determine whether the ROC curves for the 3 screens were significantly different from one another. For correlated samples (e.g., simultaneous comparison of 3 screens in the total sample), the method of DeLong et al. (1988) was used.[24] The area under the curve (AUC) statistic, a summary index of the relationship between sensitivity and specificity across a screen's range of cut scores, was also computed for each screen. An AUC of 1.0 indicates perfect ability to discriminate those with and without the disorder, and an AUC of 0.50 indicates discrimination at a level no better than chance. The optimal cut point for each screen was defined as the point closest to the upper left-hand corner of the ROC plot. To determine the statistical differences between subsequent cutoff scores, we calculated the 95% confidence intervals for each value of sensitivity, specificity, PPV, and NPV, and classified them as statistically different if the 95% confidence intervals did not overlap. The relative performance of each screen was compared across each of the gender and ethnic groups, as well as within each gender and ethnic subgroup.

RESULTS

Of the 401 participants for whom we had data, 358 (89%) with complete data reported some alcohol use in the past year and were included in the analyses. In the group of 358, the mean age was 20.6 years (standard deviation [SD], 2.1 years; range, 15–24 years) and 55% were male. While 46% were white, 49% were African-American, and 5% were of other ethnic groups.

Prevalence of DSM-IV-defined Alcohol Use Disorders

Among all subjects recruited into the study at the STD clinic ($n=401$), 29% had a current (i.e., past year) AUD. Among participants who reported some alcohol use in the past year ($n=358$), 33% had a current AUD, with 24% meeting the criteria for alcohol abuse and the remaining 9% meeting the criteria for alcohol dependence. Males were as likely as females to have an AUD (35% vs 31%; $P=.39$). However, whites were significantly more likely than blacks to have an AUD (41% vs 25%; $P<.01$).

AUDIT, CRAFFT, and CAGE Scores

In the sample of 358 the AUDIT scores ranged from 1 to 30, with an average of 8.69 (SD, 6.2); CRAFFT scores ranged from 0 to 6, with an average of 2.96 (SD, 1.8); and CAGE scores ranged from 0 to 4, with an average of 0.91 (SD, 1.1). Scores on the 3 screens were moderately correlated ($P<.01$), with r values as follows: AUDIT with CRAFFT, .71; AUDIT with CAGE, .66; and CRAFFT with CAGE, .60.

Optimal Cut Scores for Each Screen

Table 1 presents data on sensitivity, specificity, PPV, and NPV for various cut scores on the AUDIT, CRAFFT, and CAGE screens. Performance was optimal at a cut score of 9 on the AUDIT, at a cut score of 2 on the CRAFFT, and at a cut score of 1 on the CAGE. For each measure, lower cut points were associated with higher sensitivity but lower specificity.

Relative Performance of AUDIT, CRAFFT, and CAGE

Receiver operating characteristic analyses indicated that the AUDIT had the highest AUC (AUC, 0.84; standard error [SE], 0.02), followed by the CRAFFT (AUC, 0.79; SE, 0.02) and the CAGE (AUC, 0.70; SE, 0.03). All 3 screens performed at a greater than chance level ($P<.001$). Simultaneous contrast of the ROC curves for each screen, depicted in Figure 1, indicated that the 3 AUC values differed significantly from one another ($P<.001$). Paired contrasts indicated that the AUDIT had better overall performance than the CRAFFT ($P<.05$) or CAGE ($P<.001$). The CRAFFT also had better overall performance than the CAGE ($P<.001$).

Receiver Operating Characteristic Analyses by Gender

There was no significant gender-based difference in overall performance (i.e., AUC) for any of the screens (Table 2). Within the subsample of females, simultaneous contrast of the 3 screens indicated that the screens were significantly different from one another ($\chi^2=21.02$, $df=2$; $P<.001$). In females, paired contrasts revealed that both the AUDIT and CRAFFT had better overall performance than the CAGE ($P<.01$ and $P<.05$, respectively), but that the overall performance of the

Table 1. Cut Scores Reflecting the Sensitivity, Specificity, Positive Predictive Value, and Negative Predictive Value of the AUDIT, CRAFFT, and CAGE Measures for Detecting a DSM-IV-diagnosed Alcohol Use Disorder in a Sample of 358 Adolescents and Young Adults

Score AUDIT	Sensitivity	(95% CI)	Specificity	(95% CI)	PPV	(95% CI)	NPV	(95% CI)
1	1.00	(0.97 to 1.00)	0.00	(0.00 to 0.02)	0.33	(0.28 to 0.38)	–	*
2	0.99	(0.95 to 1.00)	0.08	(0.05 to 0.12)	0.35	(0.30 to 0.40)	0.95	(0.75 to 1.00)
3	0.98	(0.94 to 1.00)	0.17	(0.12 to 0.22)	0.37	(0.32 to 0.43)	0.95	(0.84 to 1.00)
4	0.96	(0.90 to 0.99)	0.26	(0.20 to 0.32)	0.39	(0.33 to 0.45)	0.95	(0.83 to 0.97)
5	0.92	(0.85 to 0.96)	0.38	(0.31 to 0.44)	0.42	(0.36 to 0.49)	0.90	(0.82 to 0.95)
6	0.90	(0.83 to 0.95)	0.51	(0.44 to 0.57)	0.48	(0.41 to 0.54)	0.91	(0.85 to 0.95)
7	0.87	(0.80 to 0.93)	0.62	(0.55 to 0.68)	0.53	(0.46 to 0.60)	0.91	(0.85 to 0.95)
8	0.82	(0.74 to 0.89)	0.72	(0.65 to 0.77)	0.59	(0.51 to 0.67)	0.89	(0.84 to 0.93)
9	0.76	(0.68 to 0.84)	0.79	(0.73 to 0.84)	0.64	(0.56 to 0.72)	0.87	(0.82 to 0.91)
10	0.70	(0.61 to 0.78)	0.83	(0.77 to 0.87)	0.67	(0.58 to 0.75)	0.85	(0.79 to 0.89)
11	0.60	(0.51 to 0.69)	0.87	(0.82 to 0.91)	0.70	(0.60 to 0.79)	0.82	(0.76 to 0.86)
CRAFFT								
1	1.00	(0.97 to 1.00)	0.11	(0.00 to 0.02)	0.36	(0.28 to 0.38)	1.00	*
2	0.94	(0.97 to 1.00)	0.33	(0.08 to 0.16)	0.41	(0.31 to 0.41)	0.92	(0.87 to 1.00)
3	0.82	(0.88 to 0.98)	0.58	(0.27 to 0.39)	0.49	(0.35 to 0.47)	0.86	(0.84 to 0.97)
4	0.71	(0.73 to 0.88)	0.75	(0.52 to 0.64)	0.59	(0.42 to 0.56)	0.84	(0.80 to 0.91)
5	0.48	(0.62 to 0.79)	0.89	(0.69 to 0.80)	0.69	(0.50 to 0.67)	0.77	(0.78 to 0.89)
6	0.24	(0.39 to 0.57)	0.98	(0.84 to 0.93)	0.85	(0.58 to 0.78)	0.72	(0.72 to 0.82)
CAGE								
1	0.69	(0.97 to 1.00)	0.63	(0.00 to 0.02)	0.48	(0.28 to 0.38)	0.80	*
2	0.50	(0.60 to 0.77)	0.82	(0.56 to 0.69)	0.58	(0.40 to 0.56)	0.77	(0.74 to 0.86)
3	0.26	(0.41 to 0.60)	0.94	(0.76 to 0.86)	0.70	(0.48 to 0.67)	0.72	(0.71 to 0.82)
4	0.02	(0.18 to 0.35)	0.99	(0.91 to 0.97)	0.50	(0.55 to 0.83)	0.67	(0.67 to 0.77)

**Confidence interval is indeterminate.*

PPV, positive predictive value; NPV, negative predictive value; AUDIT, Alcohol Use Disorders Identification Test; DSM-IV, Diagnostic and Statistical Manual of Mental Disorders, fourth edition; CI, confidence interval.

AUDIT did not differ significantly from that of the CRAFFT. Results were similar within the subsample of males.

Receiver Operating Characteristic Analyses by Race

Table 2 presents AUC values for each of the 3 screens by race. Although the overall performance of the CRAFFT and CAGE did not differ by race, the AUDIT showed better overall performance in whites than in blacks ($P<.05$). Within the subgroup of whites, simultaneous contrast of the 3 screens indicated that the screens were significantly different from one another ($P<.001$). Paired contrasts indicated that the AUDIT had a higher AUC value than the CRAFFT ($P<.001$) or the CAGE ($P<.01$), but that the AUC values for the CRAFFT and CAGE did not differ significantly. Within the subgroup of blacks, simultaneous contrast of the 3 screens indicated that the screens were significantly different from one another ($P<.01$). Paired contrasts indicated that the AUDIT and CRAFFT had better overall performance than the CAGE ($P<.01$ for each) but that the AUDIT and CRAFFT did not differ significantly. In sum, the AUDIT performed the best of the 3 screens in both whites and blacks, although the CRAFFT and the AUDIT showed similar overall performance in identifying an AUD in blacks.

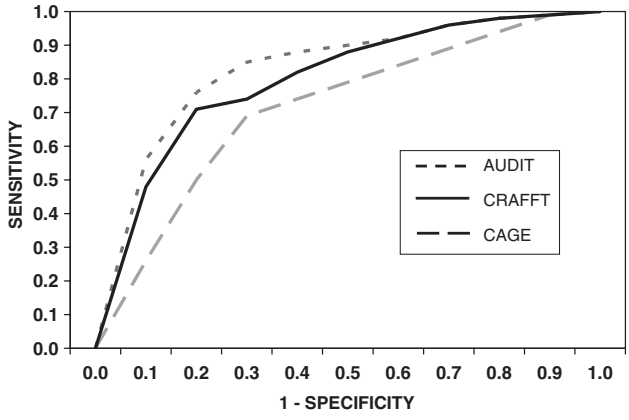

FIGURE 1. Receiver operating characteristic curves for AUDIT, CRAFFT, and CAGE screens ($n=358$).

DISCUSSION

The primary objective of our analysis was to determine which of 3 brief alcohol screens was most likely to predict the presence of an AUD in a sample of adolescents and young adults seen in an urban STD clinic. The results indicate that the 10-item AUDIT performed best overall, although the CRAFFT performed nearly as well and had excellent sensitivity. Our findings suggest that the CAGE, a widely used brief alcohol screen, has limited utility for screening adolescents and young adults in a medical setting.

What is the best screening instrument for adolescents and young adults? The ideal screening instrument should be brief, easy to remember, and easy to administer, and should show validity and reliability. Furthermore, it should be chosen on the basis of whether the clinician wishes to identify a specific AUD (alcohol abuse or alcohol dependence) or an alcohol-related drinking problem. Either the AUDIT or CRAFFT seems reasonable for this population.

The AUDIT has several advantages, including its overall good test performance, its excellent reliability and validity in a variety of populations, and its inclusion of clinically useful measures of the quantity and frequency of alcohol use.[25] Our results are consistent with the results of other studies in adolescent and young adult samples, in which the AUDIT generally showed better overall performance than the CAGE in discriminating individuals with and without an AUD.[10–12]

Table 2. Area Under the Curve Values (Standard Error) for AUDIT, CRAFFT, and CAGE Measures, Based on the Total Sample of Adolescents and Young Adults and on Gender and Race Subsamples

	Total Sample (*N*=358)	Gender		Race	
		Males (*n*=197)	Females (*n*=161)	Whites (*n*=165)	African Americans (*n*=177)
AUDIT	0.84 (0.02)	0.79 (0.03)	0.88 (0.03)	0.88 (0.03)	0.76 (0.04)
CRAFFT	0.79 (0.02)	0.75 (0.04)	0.83 (0.04)	0.78 (0.04)	0.78 (0.04)
CAGE	0.70 (0.03)	0.66 (0.04)	0.73 (0.04)	0.66 (0.05)	0.72 (0.04)

AUDIT, Alcohol Use Disorders Identification Test.

Results from our study of young persons and results from studies that involved adults generally support the AUDIT's lack of bias with regard to gender[15,25,26] and its utility across a relatively broad age range.[10–12,27] The optimal AUDIT cut score may differ depending on the age group or developmental period; the optimal score in our sample was 9, whereas the typical recommended score is 8 when used in adult samples.[8] Despite its advantages, the AUDIT's 10 items typically require a written format, and scoring is somewhat cumbersome. Therefore, the AUDIT's relative length and lack of convenience may limit its utility in certain medical settings.

Among the advantages of the CRAFFT are its brevity, its ease of administration and scoring, and its inclusion of several items that are particularly relevant to younger persons and that may need to be addressed regardless of the presence of an AUD. For example, the CRAFFT asks, "Have you ridden in a *C*ar driven by someone (including yourself) who had been drinking?" A positive response to this question should lead to a specific counseling message regardless of whether an AUD is present, because motor vehicle crashes remain the highest cause of death in individuals between the ages of 15 and 20 years. The CRAFFT may also be used to assess both alcohol and drug use with the same items, whereas the AUDIT is designed to assess alcohol only. Often, clinicians seek an initial screening test that is simple and sensitive, and our study indicated that a score of 2 on the CRAFFT had a sensitivity of 0.94 for an AUD. However, at this cut point, the specificity of CRAFFT was much lower than it was in a different study of CRAFFT at the same cut point (0.33 vs 0.81).[12] With a low specificity, low PPV, but high NPV, the CRAFFT at this cut point works best as either an initial alcohol screener to identify potentially problematic alcohol use or to rule out problematic alcohol use in this population.

The CAGE performed the worst in our sample, and the CAGE total score was neither sensitive nor specific for a current AUD. While the CAGE items may be appropriate for older adults and may have the advantage of brevity and clinician familiarity, clinicians should be aware that the CAGE consistently lacks sufficient reliability and validity in younger persons and should not be used in this population.[12]

Our results differ somewhat from those of Knight et al., who compared the AUDIT, CAGE, and CRAFFT in a primary care sample of teenagers aged 14–18.[12] In that population, CRAFFT had a relatively lower sensitivity (0.7) but excellent specificity at a cutoff of 2, whereas the AUDIT was relatively insensitive (0.54) at a cutoff of 8. Contrasts in the study findings might be explained by differences in study sample (age range 15–24 years vs 14–18 years), setting (STD clinic vs primary care setting), or prevalence of AUDs (33% vs 8%). Although sensitivity does not typically vary with disease prevalence, some have argued that large differences in prevalence could account for some of the difference in results of alcohol screening measures across populations.[12,28]

To our knowledge, previous studies have not examined differences in the performance of brief alcohol screens by gender or by race in young persons. We did not identify any gender-based differences in the overall performance of the 3 brief screens in our young sample. Although in our sample the CRAFFT and CAGE showed no significant race-based differences, the AUDIT performed slightly better in whites than in blacks. Other studies have reported small race-based differences in the AUDIT administered to adult medical patients.[15,29] These differences in results may be due to inherent cultural differences in use patterns and types of problem experienced or interpretations of AUDIT items.

Although our study is unique in its setting (STD clinic) and age range (15–24), these issues also affect study generalizability. Reporting bias is another limitation inherent to studies of alcohol and drug use, although several studies have supported the general reliability and validity of the screening measures,[29] and a clinical diagnostic assessment with good reliability and validity was used as the "gold standard."

In summary, our data suggest that the CRAFFT is a relatively simple and sensitive initial screener to identify problematic alcohol use among adolescent and young adults. Because of its relatively low specificity, it will need to be followed up by more specific questioning if one seeks to diagnose an alcohol use disorder such as alcohol abuse or alcohol dependence. The AUDIT has the best combination of sensitivity and specificity and will work well in settings where one can use a written screening instrument. In contrast, the commonly used CAGE questions do not appear to be an appropriate screening instrument for adolescents and young adults based on our data and that of others. The choice of an appropriate screening measure is just the first step in the assessment process. When a potential alcohol problem is detected, clinicians may wish to administer one of the brief alcohol interventions that have proved to be effective in primary care settings,[31,32] or they may opt to refer the patient for specialized counseling or treatment.[33]

This work was supported by federal grants K23AA00303-01, K02 AA000291, K-08-AA00280, and K01 AA00324 from the National Institute of Alcohol Abuse and Alcoholism.

REFERENCES

1. **Committee on Developing a Strategy to Reduce and Prevent Underage Drinking, Bonnie RJ, O'Connell ME, eds.** Reducing Underage

Drinking: A Collective Responsibility. Washington, DC: National Academies Press, 2003.
2. **U.S. Preventive Services Task Force.** Guide to Clinical Preventive Services, 2nd ed. Baltimore, MD: Williams & Wilkins; 1996.
3. **O'Connor PG, Schottenfeld RS.** Patients with alcohol problems. N Engl J Med. 1998;338:592–602.
4. **Kelly TM, Donovan JE, Cornelius JR, Bukstein OG, Delbridge TR, Clark DB.** Psychiatric disorders among older adolescents treated in emergency departments on weekends: a comparison with a matched community sample. J Stud Alcohol. 2003;64:616–22.
5. **Deitz D, Rohde M, Bertolucci M, Dufour M.** Prevalence of screening for alcohol use by physicians during routine physical examination. Alcohol Health Res World. 1994;18:162–8.
6. **Conigliaro J, Gordon AJ, McGinnis KA, Rabeneck L, Justice AC.** How harmful is hazardous alcohol use and abuse in HIV infection: do health care providers know who is at risk? J Acquir Immune Defic Syndr. 2003;33:521–5.
7. **Ewing JA.** Detecting alcoholism. The CAGE questionnaire. JAMA. 1984; 252:1905–7.
8. **Babor T, Ramon de la Fuente J, Saunders J, Grant M.** The Alcohol Use Disorders Identification Test: Guidelines for Use in Primary Health Care. WHO Publication No. 89.4. Geneva, Switzerland: World Health Organization; 1989.
9. **Knight JR, Shrier LA, Bravender TD, Farrell M, Vander Bilt J, Shaffer HJ.** A new brief screen for adolescent substance abuse. Arch Pediatr Adolesc Med. 1999;153:591–6.
10. **Chung T, Colby SM, Barnett NP, Rohsenow DJ, Spirito A, Monti PM.** Screening adolescents for problem drinking: performance of brief screens against DSM-IV alcohol diagnoses. J Stud Alcohol. 2000;61: 579–87.
11. **Clements R.** A critical evaluation of several alcohol screening instruments using the CIDI-SAM as a criterion measure. Alcohol Clin Exp Res. 1998;22:985–93.
12. **Knight JR, Sherritt L, Harris SK, Gates EC, Chang G.** Validity of brief alcohol screening tests among adolescents: a comparison of the AUDIT, POSIT, CAGE, and CRAFFT. Alcohol Clin Exp Res. 2003;27:67–73.
13. **O'Hare T, Tran TV.** Predicting problem drinking in college students: gender differences and the CAGE questionnaire. Addict Behav. 1997; 22:13–21.
14. **Cherpitel CJ.** Screening for alcohol problems in the U.S. general population: a comparison of the CAGE and TWEAK by gender, ethnicity, and services utilization. J Stud Alcohol. 1999;60:705–11.
15. **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;19:628–34.
16. **Bradley KA, Boyd-Wickizer J, Powell SH, Burman ML.** Alcohol screening questionnaires in women: a critical review. JAMA. 1998;280:166–71.
17. **Aertgeerts B, Buntinx F, Fevery J, Ansoms S.** Is there a difference between CAGE interviews and written CAGE questionnaires? Alcohol Clin Exp Res. 2000;24:733–6.
18. **Maisto SA, Conigliaro J, McNeil M, Kraemer K, Kelley ME.** An empirical investigation of the factor structure of the AUDIT. Psychol Assess. 2000;12:346–53.
19. **Chung T, Colby SM, Barnett NP, Monti PM.** Alcohol use disorders identification test: factor structure in an adolescent emergency department sample. Alcohol Clin Exp Res. 2002;26:223–31.
20. **Kelly TM, Donovan JE.** Confirmatory factor analyses of the alcohol use disorders identification test (AUDIT) among adolescents treated in emergency departments. J Stud Alcohol. 2001;62:838–42.
21. **Martin CS, Kaczynski NA, Maisto SA, Bukstein OM, Moss HB.** Patterns of DSM-IV alcohol abuse and dependence symptoms in adolescent drinkers. J Stud Alcohol. 1995;56:672–80.
22. **Martin CS, Kaczynski NA, Maisto SA, Tarter RE.** Polydrug use in adolescent drinkers with and without DSM-IV alcohol abuse and dependence. Alcohol Clin Exp Res. 1996;20:1099–108.
23. **Martin C, Pollock N, Bukstein O, Lynch K.** Inter-rater reliability of the SCID alcohol and substance use disorders section among adolescents. Drug Alcohol Depend. 2000;59:173–6.
24. **DeLong ER, DeLong DM, Clarke-Pearson DL.** Comparing the areas under two or more correlated receiver operating characteristic curves: a non parametric approach. Biometics. 1998;44:837–45.
25. **Reinert DF, Allen JP.** The Alcohol Use Disorders Identification Test (AUDIT): a review of recent research. Alcohol Clin Exp Res. 2002;26: 272–9.
26. **Steinbauer JR, Cantor SB, Holzer CE III, Volk RJ.** Ethnic and sex bias in primary care screening tests for alcohol use disorders. Ann Intern Med. 1998;129:353–62.
27. **Fleming MF, Barry KL, MacDonald R.** The alcohol use disorders identification test (AUDIT) in a college sample. Int J Addict. 1991;26: 1173–85.
28. **Kraemer HC.** Evaluating Medical Tests: Objective and Quantitative Guidelines. Newbury Park, CA: Sage Publications; 2000.
29. **Cherpitel CJ.** Differences in performance of screening instruments for problem drinking among blacks, whites and Hispanics in an emergency room population. J Stud Alcohol. 1998;59:420–6.
30. **Clark DB, Winters KC.** Measuring risks and outcomes in substance use disorders prevention research. J Consult Clin Psychol. 2002;70:1207–23.
31. **Monti P, Barnett N, O'Leary T, Colby S.** Motivational enhancement for alcohol-involved adolescents. In: Monti P, Colby S, O'Leary T, eds. Adolescents, Alcohol and Substance Abuse. New York, NY: Guilford Press; 2001:145–82.
32. **Fleming MF, Barry KL, Manwell LB, Johnson K, London R.** Brief physician advice for problem alcohol drinkers. A randomized controlled trial in community-based primary care practices. JAMA. 1997;277: 1039–45.
33. **Samet JH, Rollnick S, Barnes H.** Beyond CAGE. A brief clinical approach after detection of substance abuse. Arch Intern Med. 1996; 156:2287–93.