ORIGINAL INVESTIGATION

Screening for Alcohol Problems in Primary Care

A Systematic Review

David A. Fiellin, MD; M. Carrington Reid, PhD, MD; Patrick G. O'Connor, MD, MPH

Background: Primary care physicians can play a unique role in recognizing and treating patients with alcohol problems.

Objective: To evaluate the accuracy of screening methods for alcohol problems in primary care.

Methods: We performed a search of MEDLINE for years 1966 through 1998. We included studies that were in English, were performed in primary care, and reported the performance characteristics of screening methods for alcohol problems against a criterion standard. Two reviewers appraised all articles for methodological content and results.

Results: Thirty-eight studies were identified. Eleven screened for at-risk, hazardous, or harmful drinking; 27 screened for alcohol abuse and dependence. A variety of screening methods were evaluated. The Alcohol Use Disorders Identification Test (AUDIT) was most effective in identifying subjects with at-risk, hazardous, or harmful drinking (sensitivity, 51%-97%; specificity, 78%-96%), while the CAGE questions proved superior for detecting alcohol abuse and dependence (sensitivity, 43%-94%; specificity, 70%-97%). These 2 formal screening instruments consistently performed better than other methods, including quantity-frequency questions. The studies inconsistently adhered to methodological standards for diagnostic test research: 3 (8%) provided a full description of patient spectrum (demographics and comorbidity), 30 (79%) avoided workup bias, 12 (of 34 studies [35%]) avoided review bias, and 21 (55%) performed an analysis in pertinent clinical subgroups.

Conclusions: Despite methodological limitations, the literature supports the use of formal screening instruments over other clinical measures to increase the recognition of alcohol problems in primary care. Future research in this field will benefit from increased adherence to methodological standards for diagnostic tests.

Arch Intern Med. 2000;160:1977-1989

From the Department of Internal Medicine, Yale University School of Medicine, New Haven, Conn (Drs Fiellin, Reid, and O'Connor); and the Clinical Epidemiology Unit, Veterans Affairs Connecticut Healthcare System, West Haven (Dr Reid).

EXCESSIVE alcohol consumption is associated with considerable morbidity and mortality and substantial direct and indirect economic costs.[1] It is estimated that alcohol use is responsible for 100 000 deaths annually and a $100 billion cost.[1] Primary care physicians provide routine care for a large number of patients with alcohol problems; prevalence rates range from 2% to 29%, depending on the type of disorder, in ambulatory patients.[2-4]

Primary care physicians are encouraged by the National Institute on Alcohol Abuse and Alcoholism to screen patients not only for alcohol abuse and dependence, but also for alcohol consumption that would place them at risk for current or future adverse health events.[5,6] The rationale for this recommendation is that primary care physicians can play an instrumental role in recognizing alcohol problems, initiating therapy, providing advice for further treatment options, monitoring response to therapy, and promoting relapse prevention.[7,8]

Studies of screening instruments in primary care have focused on a wide spectrum of alcohol consumption, including at-risk, heavy, or harmful drinking, and alcohol abuse and dependence (**Table 1**). At-risk or hazardous drinking is usually defined by establishing a threshold amount of alcohol consumption (eg, daily, weekly, or per occasion) and is also referred to as *problem*, *heavy*, or *excessive drinking*.[5] This pattern of drinking is thought to put patients at risk for alcohol-related consequences either because of the amount they drink or because of the effect of alcohol on comorbid medical conditions. Harmful drinkers exhibit physical or psychological harm from alcohol consumption but may not meet criteria for alcohol dependence.[5,9] Patients with alcohol abuse

MATERIALS AND METHODS

We searched the MEDLINE database by using specific medical subject heading and text words to identify candidate articles for review (**Table 2**). Potential articles were examined to determine if they met the following eligibility criteria: (1) were published in peer-reviewed journals between 1966 and 1998, (2) were written in English, (3) were performed in a primary care setting, (4) examined the performance characteristics of screening methods for alcohol problems, (5) compared a screening method to a criterion standard, and (6) reported performance characteristics (eg, sensitivity and specificity) for the method.

Evaluating the accuracy of a screening instrument requires that a reference or criterion standard be used to determine whether a diagnosis is present or absent. The choice of criterion standard depends on the disorder that is the target of the screening. While accurate diagnosis of an alcohol problem can be difficult because of the complexity of the disorders and the variety of diagnostic schemas available (Table 1), standardized diagnostic instruments exist. For the purpose of this review, we considered that a study compared a screening method with a criterion standard if an identified diagnostic instrument, eg, Structured Clinical Interview for the *Diagnostic and Statistical Manual of Mental Disorders, Revised Third Edition*,[11] or operational definition (eg, quantity and frequency of alcohol consumption) was used to establish the presence or absence of an alcohol problem.

The goal in performing this systematic review was to evaluate the evidence, and the quality of that evidence, for screening instruments for alcohol problems in primary care. To evaluate the quality of the evidence, we determined whether the methods used in the studies conformed to standards designed to increase validity in this type of research. We appraised each report according to standards that are used to assess the quality of evidence in screening and diagnostic test research.[12-15] All eligible articles were appraised by means of a standardized form to record pertinent study characteristics and results according to prespecified coding criteria. The methodological standards and the realms evaluated are described below.

DESCRIPTION OF PATIENT SPECTRUM

An adequate description of the spectrum of patients included in a study can help clinicians know whether to generalize the results to their patients. To allow clinicians to decide whether the study populations were representative of an unbiased selection of patients, and similar to their own, we considered the spectrum of the patients included in the studies. A study met this standard if the following information was provided about the study population: (1) demographics (age and sex distribution), (2) comorbidity (medical and psychiatric), and (3) eligibility criteria and the number of eligible and screened subjects (ie, participation rate).

AVOIDANCE OF WORKUP BIAS

Workup bias[12] occurs when subjects with a positive (or negative) result on a screening test preferentially receive the criterion standard evaluation and can distort a test's performance. For instance, if patients with positive, as opposed to negative, results on screening tests preferentially receive the criterion standard evaluation, the sensitivity of the test can be falsely elevated because of the incorrect exclusion of subjects (false negatives) from the analysis. Therefore, a study met this standard if all subjects received both the screening and criterion standard test.

AVOIDANCE OF REVIEW BIAS

Review bias[12,15] occurs when knowledge of a subject's results on a screening examination affects the interpretation of the criterion standard test. This can occur when the screening test and criterion standard procedure are not performed in a blinded fashion. For instance, a patient's response on a screening evaluation (eg, 4 positive responses to the CAGE questions, a questionnaire for detecting alcohol problems ["Have you ever felt you should *c*ut down on your drinking?" "Have people *a*nnoyed you by criticizing your drinking?" "Have you ever felt bad or *g*uilty about drinking?" "Have you ever taken a drink first thing in the morning (*e*ye-opener) to steady your nerves or get rid of a hangover?"]) could potentially influence scoring on a subsequent diagnostic interview to evaluate for alcohol use disorders. Failure to meet this standard could result in an overestimation of the test's performance. To assess for avoidance of review bias, we considered the sequence of the screening and criterion standard evaluation and whether blinding was described. To meet this standard, we required that investigators report that blinding was performed.

ANALYSIS OF PERTINENT CLINICAL SUBGROUPS

Whereas studies may evaluate the accuracy of a screening test in a population with a broad range of drinking disorders and patient characteristics, clinicians may be interested in a test's accuracy in a particular clinical subgroup. For instance, test accuracy may vary according to demographic or clinical (eg, severity of alcohol problem) factors. If study results are presented as an aggregate, the clinician can only extrapolate results from one group to another without any assurance that the test performs equally well in each group. To check for the clinical utility of the results, we determined whether an analysis was performed on pertinent clinical subgroups. We considered that this standard was met if there was a separate analysis by demographic characteristic or diagnostic category, eg, current vs lifetime disorder.

and dependence experience marked and repeated negative physical and social effects from alcohol.[10] These diagnostic classification schemes can be used by clinicians to stratify patients with respect to severity, prognosis, and appropriate treatment regimens. In addition, they should be considered when a strategy is chosen for identifying patients with alcohol problems. For instance, a method that performs well in detecting patients with alcohol abuse or dependence may perform poorly in identifying patients drinking at harmful or hazardous levels.

The goal of the current review is to answer the clinical question: "Are there effective screening strategies to

Table 1. Categories and Definitions for Patterns of Alcohol Use*

Category	Organization	Definition
Moderate drinking	NIAAA	Men, ≤2 drinks/d; women, ≤1 drink/d; >65 y, ≤1 drink/d
At-risk drinking	NIAAA	Men, >14 drinks/wk or >4 drinks/occasion; women, >7 drinks/wk or >3 drinks/occasion
Hazardous drinking	WHO	At risk for adverse consequences from alcohol
Harmful drinking	WHO	Alcohol is causing physical or psychological harm
Alcohol abuse	APA	≥1 of the following events in a year; recurrent use resulting in failure to fulfill major role obligations; recurrent use in hazardous situations; recurrent alcohol-related legal problems (eg, DUI); continued use despite social or interpersonal problems caused or exacerbated by alcohol
Alcohol dependence	APA	≥3 of the following events in a year: tolerance; increased amounts to achieve effect; diminished effect from same amount; withdrawal; a great deal of time spent obtaining alcohol, using it, or recovering from its effect; important activities given up or reduced because of alcohol; drinking more or longer than intended; persistent desire or unsuccessful efforts to cut down or control alcohol use; use continued despite knowledge of having a psychological problem caused or exacerbated by alcohol

**NIAAA indicates National Institute on Alcohol Abuse and Alcoholism; WHO, World Health Organization; APA, American Psychiatric Association; and DUI, driving under the influence.*

Table 2. MeSH Terms and Text Words Used to Identify Articles on Screening for Alcohol Problems in Primary Care*

Concept	MeSH Terms	Text Words
Alcohol problems	Alcoholism, alcohol drinking, ethanol	Alcohol abuse, alcohol dependence, alcohol use disorders, problem drinking, hazardous drinking, heavy drinking,
Detection	Mass screening, substance abuse detection, questionnaires, sensitivity and specificity, ROC curves, predictive value of tests	Screening, CAGE, MAST (BMAST, SMAST) AUDIT, TWEAK, T-ACE, GGT, AST, ALT, SGPT, SGOT, carbohydrate-deficient transferrin
Primary care	Primary health care; outpatient clinics, hospital; ambulatory care; ambulatory care facilities; family practice	Primary care, general medicine

**MeSH indicates medical subject heading; ROC, receiver operating characteristic; CAGE, a questionnaire for alcoholism evaluation ("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 drinking?" "Have you ever taken a drink first thing in the morning [eye-opener] to steady your nerves or get rid of a hangover?"); MAST, Michigan Alcoholism Screening Test; BMAST, Brief MAST; SMAST, Short MAST; AUDIT, Alcohol Use Disorders Identification Test; TWEAK (tolerance, worry, eye-opener, amnesia, kut down), a combination of items from CAGE and MAST; T-ACE, self-administered screening questionnaire that asks about tolerance to alcohol, being annoyed by others' comments about drinking, attempts to cut down, and having a drink first thing in the morning ("eye-opener"); GGT, γ-glutamyltransferase; AST, aspartate aminotransferase; ALT, alanine aminotransferase; SGPT, serum glutamic-pyruvic transaminase; and SGOT, serum glutamic-oxaloacetic transaminase.*

identify patients with alcohol problems in primary care settings?" To answer this question, we reviewed the literature on the detection of alcohol problems in primary care settings and assessed the strength of the evidence in support of these efforts, on the basis of an appraisal of the methods used in these studies.

RESULTS

Our MEDLINE database search yielded 373 citations. We excluded nonresearch publications such as reviews, letters, and editorials (n=56); studies that were not performed in primary care settings (n=73); studies that did not examine the performance characteristics of screening methods for alcohol problems (n=151); and studies that did not compare a screening method with a criterion standard (n=55), leaving 38 articles in the final sample. Some studies evaluated more than 1 screening instrument. The number of studies that examined the performance of each screening test is as follows: the Alcohol Use Disorders Identification Test (AUDIT) or a variation (n=9), the CAGE questions or a variation (n=15), the Michigan Alcoholism Screening Test (MAST) or a variation (n=8), the 2-question screen proposed by Cyr and Wartman[16] (n=3), mental or general health screens (n=4), quantity-frequency questions (n=6), and clinical indicators such as clinician recognition or laboratory tests (n=7).

The main focus of screening was at-risk, heavy, and harmful drinking in 11 of the studies and alcohol abuse or dependence in 27. We will discuss studies that evaluated screening tests for at-risk, heavy, and harmful drinking, followed by studies that examined tests for detecting alcohol abuse and dependence. Articles that report results of screening for at-risk, heavy, or harmful drinking, as well as alcohol abuse and dependence, are discussed in the first section.

AT-RISK, HEAVY, AND HARMFUL DRINKING

Description of Studies

The 11 studies on screening for at-risk, heavy, and harmful drinking were performed in a variety of primary care settings (**Table 3**). Five of the studies were performed outside of the United States.[3,17-20] The mean age of the subjects, when reported, ranged from 35 to 47 years,[2,3,17,18,20] while 1 study[4] included only subjects aged 60 years or older. Between 30% and 100% of the subjects were male. The prevalence of alcohol problems in the populations ranged from 1% to 44% and varied by sex and disorder. Finally, in all of the studies, either the

Table 3. Screening for At-Risk, Harmful, or Hazardous Alcohol Consumption[a]

Source, y	Setting	Screening Goal	Screening Instrument	Criterion Standard
Saunders et al,[17] 1993	Community physicians' offices, hospital-based clinics, community health centers	Hazardous Harmful	AUDIT ≥8	Hazardous[b] Harmful[c]
Piccinelli et al,[3] 1997	Community practice	Hazardous Abuse Dependence	AUDIT ≥5	*ICD-10* Hazardous[f]
Volk et al,[2] 1997	Academic-affiliated family practice center	At-risk drinking[g] Abuse Dependence	AUDIT ≥8	AUDADIS *ICD-10*
Bush et al,[21] 1998	Veterans Affairs general medicine clinics	Heavy drinking[h] Abuse Dependence	AUDIT ≥8 AUDIT-C AUDIT-3	Trilevel WHO interview DIS
Bradley et al,[22] 1998	Veterans Affairs general medicine clinics	Heavy drinking[i]	AUDIT alcohol consumption questions (first 3 questions)	Trilevel WHO interview
Bradley et al,[23] 1998	Veterans Affairs general medicine clinics	Heavy drinking[k] Abuse Dependence	AUDIT ≥8 CAGE ≥2 Augmented CAGE[l] ≥2	Trilevel WHO interview DIS
King,[18] 1986 Adams et al,[4] 1996	Academic-affiliated health center Community practices	At-risk drinking Excessive drinking	CAGE ≥2 CAGE ≥2 Quantity-frequency questionnaire	At-risk drinking[m] Excessive drinking[n]
Taj et al,[24] 1998	Academic-affiliated family practice clinic	At-risk drinking Abuse Dependence	Single-question screen	At-risk drinking[o] CIDI
Aithal et al,[19] 1998	Outpatient medical clinic	Heavy drinking	CAGE ≥2 CDT[q] Laboratory values	Heavy drinking[p]
Sillanaukee et al,[20] 1998	Community practices	Heavy drinking	CDT Laboratory values	Heavy drinking[r]

[a] *AUDIT indicates Alcohol Use Disorders Identification Test;* ICD-10, International Classification of Diseases, 10th Revision (ICD-10)*; AUDADIS, Alcohol Use Disorders and Associated Disabilities Interview Schedule; NS, not specified; AUDIT-C, the first 3 questions of the AUDIT ("How often did you have a drink containing alcohol in the past year?," "How many drinks did you have on a typical day when you were drinking in the past year?," and "How often did you have 6 or more drinks on 1 occasion in the past year?"); AUDIT-3, the third question of the AUDIT ("How often did you have 6 or more drinks on 1 occasion in the past year?"); WHO, World Health Organization; DIS, Diagnostic Interview Schedule; CAGE, a questionnaire for alcoholism evaluation (*C*, Have you ever felt the need to cut down on your drinking?* A*, Have you ever felt annoyed by criticism of your drinking?* G*, Have you ever felt guilty about your drinking?* E*, Have you ever taken a drink [eye-opener] first thing in the morning?); CIDI, Composite International Diagnostic Interview; CDT, carbohydrate-deficient transferrin; MCV, mean corpuscular volume; GGT, γ-glutamyltransferase; AST, aspartate aminotransferase; and ALT, alanine aminotransferase.*

[b] *Men, 5 or more drinks per day; women, 3 or more drinks per day.*

[c] *Medical, trauma, domestic, or social problems caused by alcohol.*

[d] *Range of means.*

[e] *Data from Saunders et al.[25]*

[f] *Men, 3 to 7 drinks almost every day or 7 or more drinks at least 3 times a week; women, 2 to 5 drinks almost every day or 5 or more drinks at least 3 times a week.*

[g] *One negative consequence of alcohol use in past year or 40 g or more (1 drink is equivalent to 12 g of alcohol) per day in men, 20 g or more per day in women, or 5 or more drinks on 1 occasion in past month.*

Age, y	Sex, % M	Prevalence	Sensitivity	Specificity
35-39[d]	55[e]	Hazardous: 9% Harmful: 11%	Hazardous: 97% Harmful: 95%	Hazardous: 78% Harmful: 85%
42	36	Hazardous: 13% Harmful: 3% Dependence: 2%	Harmful or hazardous or dependence: 84%	Harmful or hazardous or dependence: 90%
39-47[d]	NS	Hazardous: <9% Harmful: 1% Abuse or dependence: male, 11%-14%; female, 5%-7%	At-risk drinking or dependence: 51%	At-risk drinking or dependence: 96%
83% of subjects (≥60)	100	Heavy: 35% Abuse or dependence: 21% Heavy or abuse or dependence: 41%	Heavy: AUDIT, 59%; AUDIT-C, 98%; AUDIT-3, 79% Heavy or abuse or dependence: AUDIT, 58%; AUDIT-C, 95%; AUDIT-3, 77% Current abuse or dependence: AUDIT, 71%; AUDIT-C, 90%; AUDIT-3, 81%	Heavy: AUDIT, 91%; AUDIT-C, 57%; AUDIT-3, 79% Heavy or abuse or dependence: AUDIT, 95%; AUDIT-C, 60%; AUDIT-3, 83% Current abuse or dependence: AUDIT, 85%; AUDIT-C, 45%; AUDIT-3, 69%
67	100	Heavy: 15%-21%[j]	>4 drinks/d: 79% 6 drinks/occasion: 69% 14 drinks/wk: 54%	>4 drinks/d: 90% 6 drinks/occasion: 91% 14 drinks/wk: 93%
87% of subjects (>60)	100	Heavy: 34%[j]	Heavy: AUDIT, 57%; CAGE, 49%, augmented CAGE, 65% Heavy and/or abuse or dependence: AUDIT, 55%; CAGE, 53%; augmented CAGE, 72%	Heavy: AUDIT, 96%; CAGE, 75%; augmented CAGE, 74% Heavy and/or abuse or dependence: AUDIT, 96%; CAGE, 81%; augmented CAGE, 74%
40 (16-85)	30	5%	84%	95%
84% of subjects (60-75); 16% of subjects (>75)	44	Excessive: 15% men, 12% women Abuse: 9% men, 3% women	Excessive drinking: 14%	Excessive drinking: 97%
NS	NS	At risk, abuse/dependence: 25%	62%	93%
49 (median) (20-85)	53	16%	CAGE: 69% CDT: 69% MCV: 54% GGT: 77%	CAGE: 95% CDT: 81% MCV: 85% GGT: 81%
Men, 42 (20-60); women, 40 (20-60)	71	Men, 44%; women, 32%	CDT: 39% MCV: 28% AST: 12% ALT: 28% GGT: 33%	CDT: 29% MCV: 40% AST: 20% ALT: 29% GGT: 34%

[h] *More than 14 drinks per week or 5 or more drinks on 1 occasion in the past or a typical month based on the Trilevel alcohol consumption interview.*
[i] *More than 4 drinks per typical drinking day or 6 or more drinks on 1 occasion at least weekly or more than 14 drinks in a typical week.*
[j] *Heavier drinkers oversampled 2:1.*
[k] *More than 14 drinks per week in a typical month or 5 or more drinks in a day at least monthly.*
[l] *CAGE questions, the first 2 questions of the AUDIT, and the question, "Have you ever had a drinking problem?"*
[m] *More than 64 g of alcohol per day.*
[n] *Men, more than 14 drinks per week; women, more than 7 drinks per week.*
[o] *Men, more than 14 drinks per week or 4 drinks per occasion; women, more than 7 drinks per week or 3 drinks per occasion.*
[p] *More than 400 g of alcohol per week (1 drink is equivalent to 12 g of alcohol).*
[q] *Men, positive at greater than 21 U/L; women, positive at greater than 26 U/L.*
[r] *Men, more than 280 g of alcohol per week and/or CAGE score of 3 or more; women, more than 190 g of alcohol per week and/or CAGE score of 2 or more.*

screens were self-administered or screening was conducted by the research staff.

Accuracy of Screening Instruments

To allow for meaningful comparisons within screening instruments and across studies, this section describes the accuracy of the screening methods organized by instrument.

AUDIT. Six studies evaluated the AUDIT for detecting at-risk, harmful, or heavy drinking[2,3,17,21-23] (Table 3). The AUDIT had a sensitivity of 97% and a specificity of 78% for hazardous use and a sensitivity of 95% and a specificity of 85% for harmful use when a cutoff of 8 or more was used.[17] Using the same cutoff, but different criterion standards, others have reported sensitivities between 51% and 59% and specificities of 91% to 96% for detecting at-risk drinking or heavy drinking.[2,21-23] Piccinelli and colleagues[3] reported a sensitivity of 84% and a specificity of 90% for combined hazardous, harmful, or dependent drinking when a cutoff of 5 or more was used. A brief version of the AUDIT that includes only the first 3 (consumption) questions was evaluated and found to have a sensitivity of 54% to 98% and a specificity of 57% to 93% for various definitions of heavy drinking.[21-23]

CAGE Questionnaire. Four studies evaluated the CAGE questionnaire as a screening tool for at-risk, harmful, or hazardous drinkers in primary care.[4,18,19,23] King[18] evaluated the ability of the CAGE questions to detect at-risk drinkers, defined as those who consumed 64 g or more of alcohol per day, and found that this 4-item screen had a sensitivity of 84% and a specificity of 95% when a cutoff of 2 or more positive responses was used. Using the same criteria for a positive score, Adams et al,[4] however, found that the CAGE questionnaire had a sensitivity of 14% and a specificity of 97% for detecting at-risk drinking (according to National Institute of Alcohol Abuse and Alcoholism criteria) among patients older than 60 years. The CAGE questionnaire had a sensitivity between 49% and 69% and a specificity between 75% and 95% in screening for patients with heavy drinking.[19,23] An augmented CAGE questionnaire, which includes the 4 CAGE questions, the first 2 quantity and frequency questions of the AUDIT, and a question pertaining to history of drinking problems, had a sensitivity of 65% and a specificity of 74% in 1 study.[23]

Other Screening Methods. Three other studies examined the operating characteristics of a screen for this spectrum of drinking in primary care settings.[19,20,24] Taj et al[24] evaluated the properties of a single question, "On any single occasion during the past 3 months, have you had more than 5 drinks containing alcohol?" among primary care patients. This single-item screen had a sensitivity of 62% and a specificity of 93% for detecting problem drinkers.[24] Two studies investigated the operating characteristics of selected laboratory values for identifying patients with this spectrum of alcohol problem. Carbohydrate-deficient transferrin, a new serological marker for recent alcohol ingestion, had a sensitivity of 39% to 69% and a specificity of 29% to 81% for heavy drinking.[19,20] In addition, mean corpuscular volume, aspartate aminotransferase, alanine aminotransferase, and γ-glutamyltransferase had limited utility as screening tests for this disorder,[19,20] although 1 group found a sensitivity of 77% and a specificity of 81% for γ-glutamyltransferase.[19]

ALCOHOL ABUSE AND DEPENDENCE

Description of Studies

The 27 studies on screening for alcohol abuse and dependence are described in **Table 4**. These studies were conducted in a variety of primary care settings, with 4 of the studies performed outside of the United States.[26-29] The mean age of subjects in studies reporting demographic information ranged from 36 to 72 years. Males represented between 19% and 100% of the subjects. The prevalence of alcohol problems in the population ranged from 2% to 41%, depending on the diagnosis and whether lifetime or current criteria were applied. Finally, in most studies (66%), screening was performed by research staff, whereas in the remaining investigations the screen was either self-administered (15%) or clinician-administered (19%).

Accuracy of Screening Instruments

AUDIT. The AUDIT is designed to detect less severe alcohol problems, such as hazardous and harmful drinking, as well as alcohol abuse and dependence disorders. Five studies have examined the performance of the AUDIT as a screening tool for alcohol abuse or dependence. The operating characteristics of the screen varied with the cutoff used to determine positive results of a screen and whether one is interested in detecting a lifetime (ie, if patients met criteria for these disorders at any point in their life) or current diagnosis. For instance, in 1 study,[30] the AUDIT had a sensitivity of 61% and a specificity of 90% for a current alcohol use disorder with the use of a cutoff of 8. Changing the cutoff score to greater than 11 resulted in an expected decrease in sensitivity of (40%) and an increase in specificity (96%). The performance characteristics changed dramatically when the investigators considered lifetime alcohol use disorders. In this situation, the AUDIT had a sensitivity of 46% and 30% with a specificity of 90% and 97% with the use of cutoff scores of 8 and 11, respectively.[30] Other investigators found that the AUDIT had a sensitivity of 63% and 93% and a specificity of 96% and 96%, for a lifetime or current diagnosis, respectively, of alcohol abuse or dependence.[31] The AUDIT did not perform as well as a screening test in a study by Schmidt et al.[32] In this study, the AUDIT had a sensitivity of 38% with a specificity of 95% for a lifetime diagnosis of alcohol abuse or dependence. These results are similar to those obtained by Morton et al[33] with a cutoff of 8 in a population older than 65 years. In this study, the AUDIT had a sensitivity of 33% and a specificity of 91%.[33] The AUDIT was noted to

have different performance characteristic within different ethnic and sex populations.[34] In 1 study, the AUDIT, with a cutoff of 8 for a positive test, had a sensitivity between 70% and 92% with a specificity of 73% to 94%, with variation based on sex and ethnic background.[34]

CAGE Questionnaire. Ten studies evaluated the performance of the CAGE questionnaire in screening patients for alcohol abuse and/or dependence in primary care settings.[28,33-41] Sensitivities of 21% to 94% with specificities of 77% to 97% were found when a cutoff score of 2 or more was used.[34,35,38-41] Lowering the cutoff to 1 or more positive responses to CAGE questions resulted in a sensitivity of 60% to 71% and a specificity of 84% to 88%.[39,40] In older primary care populations, sensitivities ranged from 63% to 70% and specificities from 82% to 91% with CAGE questionnaire scores of 2 or more.[33,36] The CAGE questions had a sensitivity of 53% and a specificity of 93% when a combined target of alcohol abuse, dependence and harmful drinking was the goal of screening.[28]

One study investigated each of the 4 CAGE questions in a population screened for alcohol use disorders with the use of the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*[10] criteria as the criterion standard.[37] The proportion of subjects answering yes to a specific CAGE question varied by race, sex, and item. For instance, the question "Have you ever felt the need to *c*ut down on your drinking?" had a sensitivity of 63% and a specificity of 84%, whereas the question "Have you ever taken a drink (*e*ye opener) first thing in the morning?" had a sensitivity of 21% and a specificity of 95%. As with the AUDIT, the CAGE questions were noted to have varying performance characteristics within different ethnic and sex populations.[34]

MAST. Seven studies evaluated the MAST or variants of the MAST as screening tools for alcohol abuse and/or dependence.[28,30,33,42-44] As with the other screening tests, the operating characteristics of the MAST and its derivatives varied by cutoff score and diagnosis (ie, current or lifetime alcohol abuse or dependence disorder). For instance, unweighted scoring of the Short Michigan Alcoholism Screening Test (SMAST) with a cutoff of 2 or more points had a sensitivity of 82% and 100% with a specificity of 96% and 85% for detecting patients with lifetime and current diagnoses, respectively, of alcohol abuse and dependence.[43] Another study, using the same cutoff, found that the SMAST had a sensitivity of 48% and a specificity of 95%, although no distinction was made regarding current or lifetime disorders.[28] Weighted scoring of the SMAST typically uses a cutoff of 5 or more points and had a sensitivity of 57% and 66% and specificity of 80% and 80% for current and lifetime alcohol use disorders, respectively.[30] Others have reported sensitivities of 38% to 80% and specificities of 79% to 97% with the use of various cutoffs for the SMAST.[39,42,44] Finally, the recently developed geriatric version of the MAST had a sensitivity of 70% and a specificity of 80% when a cutoff score of 5 or more was used in a geriatric (>65 years old) primary care population.[33] A shortened, 9-item Self-administered Alcoholism Screening Test had a sensitivity of between 13% and 69% with a specificity of between 67% and 95% in different ethnic and sex groups in primary care.[34]

Other Instruments. Cyr and Wartman[16] found that the combination of a positive response to the question "Have you ever had a drinking problem?" and/or "When was your last drink?"(within 24 hours was considered a positive response) had a sensitivity of 91% and a specificity of 90% when MAST scores were the criterion standard. However, other investigators have attempted to replicate these findings in other primary care settings and found sensitivities between 48% and 53% and specificities between 76% and 93%.[45,46] Permutations of the single question "Have you ever had a drinking problem?" have had a sensitivity of 40% to 70% with a specificity ranging between 93% and 99%.[16,40,42,45,46]

The TWEAK questions (*t*olerance, *w*orry, *e*ye-opener, *a*mnesia, *k*ut down), a combination of items from the CAGE questionnaire and MAST developed to identify at-risk drinking among pregnant women, were found to have a sensitivity of 75% with a specificity of 90% in 1 study.[41]

Quantity-Frequency Questions. Three investigations evaluated quantity-frequency questions as a screen for alcohol abuse or dependence disorders. One study found a sensitivity of 47% and a specificity of 96%, with the use of MAST scores as the criterion standard, and a quantity cutoff score of 4 or more drinks per day.[16] Fleming and Barry[40] found sensitivities of 50% and 20% with specificities of 87% and 97%, with the use of a cutoff of 7 and 20 drinks per week, respectively. In 1 study, there was a gradual decrease in sensitivity (100%-21%) with a corresponding increase in specificity (43%-97%) as the number of drinks consumed per week increased from 0 to 24 or more.[47]

Clinical Indicators. Six studies examined clinical strategies such as clinical judgment and/or laboratory values to detect alcohol problems.[26,27,29,31,43,48] In 2 studies,[43,48] physicians identified only 36% to 77% of patients with current alcohol problems and 21% of patients with inactive alcohol problems.[48] More formal assessments have found that physician's judgment had a sensitivity of 18% to 44% with a specificity of 96% to 99% for a diagnosis of alcohol abuse and dependence.[29,31]

Attempts to formalize the use of clinical indicators have led to the creation of the Alcohol Clinical Index[26] and the use of a diagnostic grid that combines the CAGE questionnaire and features of the history and physical examination.[27] The Alcohol Clinical Index had a sensitivity of 28% and a specificity of 86% for alcohol dependence. The grid had a reported sensitivity of 99% and a specificity of 96% for alcohol dependence; however, it should be noted that the same physician provided the criterion standard diagnosis and filled out the grid.[27]

Laboratory methods for detecting alcohol problems, such as mean corpuscular volume, γ-glutamyltransferase, aspartate aminotransferase, and alanine aminotransferase, have performed poorly as screening tools.[26,43] In a receiver operating characteristic analysis, the SMAST screening test consistently outperformed physician judgment and laboratory evaluations.[43]

Table 4. Screening for Alcohol Abuse and Dependence[a]

Source, y	Setting	Screening Goal	Screening Instrument	Criterion Standard
Barry and Fleming,[30] 1993	Community clinics	Abuse; dependence	AUDIT >8; SMAST 13>5	DIS
Schmidt et al,[32] 1995	Academic-affiliated hospital-based clinic	Abuse Dependence	AUDIT >8	DIS
Isaacson et al,[31] 1994	Academic-affiliated clinic	Abuse Dependence	AUDIT >8 Clinical judgment	SCID
Steinbauer et al,[34] 1998	Family practice center, academic hospital–affiliated	Abuse Dependence	AUDIT ≥8; CAGE ≥2; SAAST ≥3	AUDADIS
Buchsbaum et al,[35] 1991	Academic-affiliated clinic	Abuse Dependence	CAGE = 2	DIS
Buchsbaum et al,[36] 1992	Academic-affiliated clinic	Problem[b]; Abuse; Dependence	CAGE = 2	DIS
Volk et al,[37] 1997	Family practice center, academic hospital–affiliated	Abuse Dependence	Individual CAGE items	AUDADIS
Chan,[38] 1994	Health care centers of county medical center	Dependence Heavy drinking[c]	CAGE ≥2	*DSM-III-R*
Cherpitel,[41] 1998	Community health clinics	Dependence	CAGE ≥2; TWEAK ≥3	CIDI
Escobar et al,[26] 1995	Health center	Dependence or men ≥60 g/d, women ≥30 g/d	CAGE Alcohol Clinical Index Laboratory values	*DSM-III*
Morton et al,[33] 1996	Veterans Affairs ambulatory care facilities	Abuse Dependence	MAST-G ≥5; CAGE ≥2; AUDIT ≥8	*DSM-III-R*
Fleming and Barry,[42] 1991	Academic-affiliated primary care clinics	Alcohol abuse Dependence	SMAST Unique questions[d]	DIS
Rumpf et al,[28] 1997	Community physicians' offices	Harmful drinking; abuse; dependence	SMAST ≥2[e]; LAST ≥2; CAGE ≥2	*DSM-III-R* or *ICD-10*
Cleary et al,[43] 1988	Academic-affiliated hospital-based clinic	Problem drinking Abuse Dependence	SMAST ≥2 Laboratory values Clinical judgment	DIS
Barry and Fleming,[44] 1990	Academic-affiliated primary care clinics	Abuse Dependence	SMAST ≥5[f] CAGE[g]	DIS
Brown and Rounds,[39] 1995	Academic-affiliated community clinic	Abuse; dependence	SMAST ≥2; CAGE ≥2	DIS
Cyr and Wartman,[16] 1988	Academic-affiliated hospital-based clinic	Alcoholism	Two-question screen[h] Quantity-frequency questions	MAST ≥5
Schorling et al,[45] 1995	Academic-affiliated outpatient clinics	Alcoholism	Two-question screen[h]	MAST ≥5
Moran et al,[46] 1990	Veterans Affairs satellite clinic	Alcoholism	Two-question screen[h]	MAST ≥5
Buchsbaum et al,[47] 1995	Academic-affiliated hospital-based clinic	Abuse; dependence	Quantity questions	DIS
Rydon et al,[29] 1992	Community physicians' offices	Alcohol-related problems[k]	Clinical judgment	SMAST ≥3 CAGE ≥2
Buchsbaum et al,[48] 1992	Academic-affiliated clinic	Problem[b]; abuse; Dependence	Physician detection	DIS
Gabrynowicz and Watts,[27] 1981	Community practice	Alcoholism[l]	Diagnostic grid	Physician judgment
Leon et al,[50] 1996	HMO primary care clinic	Dependence	SDDS-PC	*DSM-IV*
Broadhead et al,[49] 1995	Community practice and academic-affiliated practices	Abuse or dependence	SDDS-PC	*DSM-III-R*
Fleming and Barry,[40] 1991	Primary care clinics	Abuse Dependence	HSS; Quantity-frequency questions; CAGE ≥2	DIS
Hore et al,[51] 1977	Academic-affiliated general practice	Problem drinkers Alcohol addicts[m]	Spare-Time Activity Questionnaire	Psychiatrist assessment

[a] *AUDIT indicates Alcohol Use Disorders Identification Test; SMAST, Short Michigan Alcoholism Screening Test; DIS, Diagnostic Interview Schedule; NS, not specified; SCID, Structured Clinical Interview for* DSM-III-R*; CAGE, a questionnaire for alcoholism evaluation (*C*, Have you ever felt the need to* cut down *on your drinking?* A*, Have you ever felt* annoyed *by criticism of your drinking?* G*, Have you ever felt* guilty *about your drinking?* E*, Have you ever taken a drink [*eye opener*] first thing in the morning?); SAAST, Self-administered Alcoholism Screening Tests; AUDADIS, Alcohol Use Disorders and Associated Disabilities interview Schedule;* DSM-III-R, Diagnostic and Statistical Manual of Mental Disorders, Revised Third Edition*; TWEAK, a combination of items from CAGE and MAST (*tolerance, worry; eye-opener, amnesia, kut down*); CIDI, Composite International Diagnostic Interview;* DSM-III, Diagnostic and Statistical Manual of Mental Disorders, Third Edition*; MCV, mean corpuscular volume; GGT, γ-glutamyltransferase; AST, aspartate aminotransferase; ALT, alanine aminotransferase; MAST-G, geriatric MAST; LAST, Luebeck Alcohol Dependence and Abuse Screening Test;* ICD-10, International Classification of Diseases, 10th Revision *(*ICD-10*); HMO, health maintenance organization; SDDS-PC, Symptom-Driven Diagnostic System for Primary Care;* DSM-IV, Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*; and HSS, Health Screening Survey.*

[b] *One or more positive on the DIS.*

[c] *Men, 6 or more drinks per day; women, 4 or more drinks per day.*

[d] *Question 1: "Now that you have completed this form, do you think you currently have a drinking problem?" Question 2: "Thinking back, would you say that at any time in the past you had a drinking problem?"*

[e] *Original MAST items 9, 21, and 22 also diagnostic.*

Age, y, Mean (Range)	Sex, % M	Prevalence	Sensitivity	Specificity
NS	NS	Lifetime abuse or dependence: 35%	Lifetime: SMAST, 66%; AUDIT, 46% Current: SMAST, 57%; AUDIT, 61%	Lifetime: SMAST, 80%; AUDIT, 90% Current: SMAST, 80%; AUDIT, 90%
43 (19-80)	40	Lifetime abuse: 28% Dependence: 3%	38%	95%
45 (18-84)	48	Abuse or dependence: lifetime, 33%; current, 22%	Lifetime, 63%; current, 96%	Lifetime, 96%; current, 96%
43 (18-86)	30	Abuse or dependence: lifetime, 9.1%-26.5%	CAGE: 21%-69%; AUDIT: 70%-92%; SAAST: 13%-69%	CAGE: 77%-96%; AUDIT: 73%-94%; SAAST: 67%-95%
NS	30	Abuse or dependence: lifetime, 36%	74%	91%
NS	NS	Problem or abuse or dependence: 33%	70%	91%
43 (18-86)	30	15%	C: 63%; A: 24%; G: 37%; E: 21%	C: 84%; A: 93%; G: 90%; E: 95%
41	34	NS	Heavy drinking: past year, 63% Dependence: lifetime, 91%; past year, 94%	Heavy drinking: past year, 90% Dependence: lifetime, 84%; past year, 97%
41% of subjects (>50)	19	2%	Dependence, past year: CAGE ≥2: 82%; TWEAK ≥3: 75%	Dependence, past year: CAGE ≥2: 94%; TWEAK ≥3: 90%
45	57	NS	CAGE ≥1: 70%; CAGE ≥2: 24% Alcohol Clinical Index: 28% MCV: 63%; GGT: 28%; AST/ALT: 12%	CAGE ≥1: 80%; CAGE ≥2: 97% Alcohol Clinical Index: 86% MCV: 48%; GGT: 94%; AST/ALT: 91%
72	100	Abuse and dependence (current or past): 36%	MAST-G: 70%; CAGE: 63%; AUDIT: 33%	MAST-G: 80%; CAGE: 82%; AUDIT: 91%
36	35	Abuse and dependence: 30%	SMAST (weighted): 48% SMAST (unweighted): 49% Question 1: 57%; Question 2: 32%	SMAST (weighted): 86% SMAST (unweighted): 88% Question 1: 88%; Question 2: 97%
45	38	Harmful: 2.3%; abuse: 2.3%; Dependence: 8.4%	SMAST: 48%; LAST: 63%; CAGE: 53%	SMAST: 95%; LAST: 93%; CAGE: 93%
(18-75)	NS	Problem: 19% Abuse or dependence: 20%	SMAST; Problem: 68%; Abuse or dependence: lifetime, 82%; current, 100%	SMAST; Problem: 92%; Abuse or dependence: lifetime, 96% current, 85%
36	33	Abuse and dependence: 30%	45%-80%	79%-88%
39	NS	Abuse: 22%; dependence: 22%	SMAST: 38%; CAGE: 64%	SMAST: 97%; CAGE: 93%
37	37	Alcoholism[i]: 25%	91%	90%
48.3, sample 1 41.9, sample 2	26	Alcoholism[i]: 18%	53%	93%
(70-93)	100	Alcoholism[i]: 26%	48%	76%
48 (18-?)	86	Dependence: 30% of active drinkers	21%-100%[j]	43%-97%[j]
35% of subjects (17-30) 63% of subjects (31-64) 2% of subjects (>65)	44	By SMAST: 24% By CAGE: 11%	Against SMAST: 18% Against CAGE: 40%	Against SMAST: 96% Against CAGE: 97%
51	56	NS	NS	NS
NS	NS	9%	99%	96
49 (18-70)	33	Dependence: 5%	75%	97%
38 (18-70)	23	NS	38%	99%
36	35	30	HSS: 78%; ≥20 drinks/wk: 20%; ≥7 drinks/wk: 50%; CAGE: 43%	HSS: 71%; ≥20 drinks/wk: 97%; ≥7 drinks/wk: 87%; CAGE: 70%
NS	NS	41%	Problem drinking: 100% Alcohol addicts: 100%	Problem drinking: 87% Alcohol addicts: 72%

[f] *Weighted SMAST.*
[g] *See Fleming and Barry[40] for CAGE results.*
[h] *Positive responses to either or both: "Have you ever had a drinking problem?" or "When was your last drink?" (positive response is within 24 hours).*
[i] *Defined as MAST score of 5 or more.*
[j] *Sensitivities and specificities for increasing number of drinks, or standard ethanol content (SECs) per week; intervals were 0 to 2, 3 to 5, 6 to 11, 12 to 23, and 24 or more SECs.*
[k] *As measured by SMAST or CAGE.*
[l] *"A progressive, pathological reaction to alcohol, characterized by (a) detrimental effects on physical, mental, and social health and (b) partial or complete loss of the ability to abstain."*
[m] *Problem drinkers: patients whose drinking clearly created problems in their lives and who, despite this, continued drinking. Alcohol addicts: patients who showed one or more of the following: loss of control over drinking or inability to abstain, change of tolerance, and presence of withdrawal symptoms.*

Table 5. Adherence to Methodological Standards*

Source, y	Demographics Provided	Comorbidity Described	Eligibility Criteria and Participation Rate Provided	Criterion Standard Evaluation of All Screened Subjects	Blinding	Analysis of Pertinent Subgroups
Saunders et al,[17] 1993	+†	–	+	–	–	–
Piccinelli et al,[3] 1997	+	–	+	+	–	–
Volk et al,[2] 1997	–	–	+	+	–	+
Bush et al,[21] 1998	+	+	+	+	+	+
Bradley et al,[22] 1998	+	+	+	+	+	+
Bradley et al,[23] 1998	+	+	+	+	+	+
King,[18] 1986	+	–	+	–	–	–
Adams et al,[4] 1996	+	–	–	+	NI‡	+
Taj et al,[24] 1998	–	–	–	–	–	+
Aithal et al,[19] 1998	+	–	–	+	+	+
Sillanaukee et al,[20] 1998	+	–	–	+	+	+
Barry and Fleming,[30] 1993	–	–	+	+	–	+
Schmidt et al,[32] 1995	+	–	–	+	–	–
Isaacson et al,[31] 1994	+	–	+	+	–	+
Steinbauer et al,[34] 1998	+	–	+	+	–	+
Buchsbaum et al,[35] 1991	–	–	+	+	–	–
Buchsbaum et al,[36] 1992	–	–	–	+	–	+
Volk et al,[37] 1997	+	–	+	+	–	+
Chan et al,[38] 1994	+	–	–	–	–	+
Cherpitel,[41] 1998	+	–	+	+	–	–
Escobar et al,[26] 1995	+	–	–	–	+	–
Morton et al,[33] 1996	+	–	–	+	–	–
Fleming and Barry,[42] 1991	+	–	+	+	NI	–
Rumpf et al,[28] 1997	+	–	+	–	–	–
Cleary et al,[43] 1988	–	–	+	+	–	+
Barry and Fleming,[44] 1990	+	–	+	+	NI	+
Brown and Rounds,[39] 1995	–	–	–	+	–	+
Cyr and Wartman,[16] 1988	+	–	+	+	–	–
Schorling et al,[45] 1995	+	–	+	+	–	–
Moran et al,[46] 1990	+	–	–	+	+	–
Buchsbaum et al,[47] 1995	+	–	+	+	–	+
Rydon et al,[29] 1992	+	–	+	+	+	–
Buchsbaum et al,[48] 1992	+	–§	–	+	+	+
Gabrynowicz and Watts,[27] 1981	–	–	–	+	–	–
Leon et al,[50] 1996	+	–‖	+	+	+	–
Broadhead et al,[49] 1995	+	–	–	–	+	–
Fleming and Barry,[40] 1991	+	–	+	+	NI	+
Hore et al,[51] 1977	–	–	–	–	+	+
% of studies adhering to standard	76	8	61	79	35	53

*Plus indicates yes; minus, no.
†Data from Saunders et al.[25]
‡Not indicated; screening instrument was self-administered.
§Medical comorbidity described.
‖Psychiatric comorbidity described.

In another evaluation, Escobar et al[26] found that in a select group of subjects, use of the mean corpuscular volume, elevated γ-glutamyltransferase level, or aspartate aminotransferase–alanine aminotransferase ratio of 2 or more had sensitivities that ranged from 13% to 63% and specificities that ranged from 48% to 94%.

Mental and General Health Screening. Two studies evaluated a screen for mental disorders, including alcohol dependence, by means of disease-specific modules.[49,50] The alcohol items in the Symptom-Driven Diagnostic System for Primary Care cover worry about drinking, excessive drinking, and morning drinking. This screen had a sensitivity of 38% to 75% and a specificity of 97% to 99% for a current diagnosis of alcohol dependence in primary care populations.[49,50]

The Health Screening Survey, a masked screen for alcohol abuse and dependence that includes items about alcohol use buried among questions about exercise, nutrition, and smoking, was found to have a sensitivity of 78% and a specificity of 71% in a primary care population.[40] Finally, the Spare Time Activity Questionnaire, another disguised questionnaire, had a sensitivity of 100% and a specificity of 72% when compared with psychiatrist assessment of addiction to alcohol.[51]

COMPLIANCE WITH METHODOLOGICAL STANDARDS

Overall compliance with the individual standards by study is shown in **Table 5**. Most investigations (29 [76%]) provided a pertinent description of the demographic char-

Table 6. Accuracy of the AUDIT, CAGE, and SMAST by Diagnostic Category*

Instrument	Cutoff Score	At-Risk, Harmful, or Hazardous Drinking, %		Abuse/Dependence, %			
				Lifetime		Current	
		Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
AUDIT	≥8	**57**[23]	**96**[23]	33[33]	91[33]	61[30]	90[30]
		59[21]	**91**[21]	38[32]	95[32]	**66**[23]	**86**[23]
		63[38]	90[38]	**39**[23]	**89**[23]	**71**[21]	**85**[21]
		97†[17]	78[17]	46[30]	90[30]	96[31]	96[31]
		95‡[17]	85‡[17]	63[31]	96[31]		
				91[38]	84[38]		
				70-92§[34]	73-94†[34]		
CAGE	≥2	**14**[4]	**97**[4]	**43**[40]	**70**[40]	**77**[23]	**79**[23]
		49[23]	**75**[23]	**53**[23]	**86**[23]	82[41]	94[41]
		69[19]	95[19]	63[33]	82[33]	94[38]	97[38]
		84[18]	95[18]	64[39]	93[39]		
				74[35]	91[35]		
				21-69†[34]	77-96†[34]		
SMAST	≥2	68[43]	92[43]	38[39]	97[39]	100[43]	85[43]
				49[40]	**88**[40]		
				82[43]	96[43]		

*AUDIT indicates Alcohol Use Disorders Identification Test; CAGE, a questionnaire for alcoholism evaluation; and SMAST, Short Michigan Alcoholism Screening Test. Studies in boldface met standards for avoidance of workup and review bias (see the "Avoidance of Workup and Review Bias" subsections of the "Materials and Methods" section of the text).

†Hazardous use.

‡Harmful use.

§Results varied by ethnicity and sex.

acteristics (age distribution and sex of the subjects) of their respective study populations; however, only 3 (8%) described the medical or psychiatric comorbidity of the screened subjects. Overall, 23 (61%) of the articles provided eligibility criteria and rates of participation. Thirty (79%) of the investigations used methods designed to avoid workup bias, while only 12 (35%) of 34 of the studies met the standard for avoidance of review bias. Finally, 21 (55%) of the studies examined the performance of the screening instruments among different clinical subgroups.

PERFORMANCE CHARACTERISTICS OF COMMONLY USED SCREENS

Accuracy data on the most frequently investigated screening instruments, the AUDIT, CAGE questionnaire, and SMAST, are presented in **Table 6** by diagnostic category. We chose these instruments because each was evaluated in at least 3 investigations and they have been advocated for use as screening tools.[52] Of the 17 studies investigating screening for distinct diagnoses (eg, alcohol dependence) with these instruments, 6 met the 2 standards designed to minimize workup and review bias,[4,19,21,23,40,42] and 2 met all the standards.[21,23]

COMMENT

This evaluation of screening for alcohol problems in primary care reveals that a number of strategies have been evaluated in a variety of settings. To date, screening has generally been directed toward alcohol abuse and dependence. In addition, methodological standards designed to increase the validity of diagnostic test research are inconsistently adhered to in these investigations. Unfortunately, few studies have been performed with multiple instruments, allowing for a direct comparison of the screens' accuracy under similar conditions. Finally, clinicians infrequently performed screening in these investigations. Despite these limitations, the literature supports the effectiveness of select screening instruments for primary care.

On the basis of the reported accuracy of the techniques included in this evaluation, the literature supports screening for less severe alcohol problems such as at-risk, harmful, and hazardous drinking by means of the AUDIT. Designed specifically to increase detection of this spectrum of alcohol problems, and incorporating questions about quantity and frequency of consumption, there is evidence that the AUDIT has increased accuracy relative to the other screening methods, with a reported sensitivity between 57% and 97% and a specificity between 78% and 96%. Restricting the analysis to studies that reported on screening for distinct diagnoses and that also met the standards designed for avoidance of workup and review bias limits the sample to 2 studies, which reported sensitivities between 57% and 59% and specificities between 91% and 96%.[21,23] In this same population, modifications of the AUDIT restricted to consumption questions have shown promise but will need to be validated in a wider population and varied settings.[21,22]

The literature also supports screening for lifetime and current abuse or dependence disorders by means of the CAGE questions. There is evidence that the CAGE questionnaire has increased accuracy compared with other screening instruments for this spectrum of alcohol problems in primary care. This 4-item screen had a sensitivity of 43% to 94% with a specificity of 70% to 97% and performed better than the AUDIT or SMAST. Restricting the analysis to studies that reported on screening for

distinct diagnoses and that also met the standards for avoidance of workup and review bias limits the sample to 2 studies that found sensitivities of 43% to 77% and specificities of 70% to 86%.[23,40]

Our analysis provides a framework for implementing strategies to screen for alcohol problems in primary care in a number of ways. First, this research shows that, although imperfect, there are effective strategies that clinicians can use to identify unrecognized patients with alcohol problems in primary care settings. Structured instruments generally perform better than quantity-frequency questions, clinical impressions, or laboratory data that clinicians report they frequently use to detect alcohol problems in their patients.[53] Second, decisions about screening options should include a consideration of the accuracy of instruments across the spectrum of alcohol problems. For instance, the CAGE questions perform better in identifying patients with alcohol abuse and dependence. Conversely, the AUDIT is more sensitive for hazardous and harmful drinkers. This doubtlessly reflects the fact that the AUDIT includes measures of quantity and frequency that are used to establish these diagnoses. Third, the accuracy of screening instruments is responsive to clinical factors. For example, demographic characteristics and stage of diagnosis (current vs lifetime) have a profound effect on test performance. Finally, as one might expect, variation in study characteristics including patient spectrum, study methods, criterion standard, and analysis account for disparate findings regarding the accuracy of screening instruments in the medical literature. Attention to these components of study design and execution can help clinicians determine the utility of the results in their own practices.

One potential limitation of the current analysis is that it may not include all studies on screening for alcohol problems in primary care performed to date. However, we have attempted to identify the most appropriate studies in this field by using a search strategy that was both broad and unbiased. We believe that our search strategy, by using established terms, identified studies that were more likely to meet methodological standards.

A separate limitation to our conclusions is imposed by the state of the art in constructing appropriate criterion standards for the diagnosis of the various alcohol problems. The field has a number of diagnostic schemes from organizations including the World Health Organization, the National Institute on Alcohol Abuse and Alcoholism, and the American Psychiatric Association, and researchers may choose among these (and others) in selecting their diagnostic categories. In addition, there is variability in the criterion standard assessments that are used to establish an alcohol-related diagnosis. Therefore, conflicting or inconsistent results between reported accuracy of screening instruments may result from the definition used for the disorder, the choice of criterion standard, or differences between screening instruments. Further attention to developing uniform diagnostic schemes and accurate criterion standard tests will help to advance screening efforts.

Primary care clinicians should strive to identify patients across the spectrum of alcohol problems. The National Institute of Alcohol Abuse and Alcoholism recommendation that all patients who drink alcohol should be screened with the CAGE questions[6] is supported by our findings. The primary drawback to this strategy, however, is the relatively poor performance of the CAGE questions, compared with the AUDIT, in recognizing less severe drinking disorders. Therefore, in situations where time allows for more in-depth interviewing, incorporating the AUDIT may help to identify a wider spectrum of alcohol problems. Nonetheless, the concise nature of the CAGE questions makes them more amenable to primary care clinical encounters than are other longer instruments. A strategy that incorporates the CAGE questionnaire, followed by questions about quantity and frequency of consumption,[6] such as the augmented CAGE questionnaire, is pragmatic and shows promise.[23] Additional history should be obtained from all patients who have positive results with standardized screening instruments or quantity-frequency questions, and those suspected of having an alcohol disorder irrespective of their screening scores. Further diagnostic efforts to assess for specific disorders, eg, alcohol dependence or harmful or hazardous drinking, should be undertaken in this group of patients.

This review has identified substantial heterogeneity in adherence to methodological standards designed to improve validity in diagnostic test research and reporting of results. These findings undoubtedly reflect the difficulty of conducting research in clinical settings. Nonetheless, some recent studies have been successful in adhering to many or all of the standards.[21-22,50] To aid clinicians' efforts at recognizing and caring for patients with alcohol problems, future studies will benefit from increased attention to these standards. Among the areas that warrant increased attention are the reporting of characteristics of the study population, the avoidance of workup and review bias, and description of test performance in pertinent clinical subgroups. In addition, physician-based screening strategies should be empirically tested. The results in clinical settings may differ from those obtained in research because of factors that affect test performance, such as interviewing technique and competing clinical priorities, that were not accounted for in earlier studies.

Future research, by addressing these limitations, will help clinicians feel confident about extrapolating results to clinical practice and facilitate recognition of patients with alcohol problems in primary care settings. Once clinicians have identified these patients, they can begin the process of helping patients reduce the harm associated with excessive alcohol consumption.[7]

Accepted for publication January 10, 2000.

Dr Fiellin is supported by the National Institute on Drug Abuse Physician Scientist Award (NIDA K12 DA00167), Bethesda, Md. Dr Reid is supported by a Career Development Award from the Department of Veterans Affairs Health Services Research and Development Service, Washington, DC.

Reprints: David A. Fiellin, MD, Yale University School of Medicine, 333 Cedar St, PO Box 208025, New Haven, CT 06520-8025 (e-mail: david.fiellin@yale.edu).

REFERENCES

1. US Department of Health and Human Services. *Eighth Special Report to the US Congress on Alcohol and Health From the Secretary of Health and Human Services, September 1993*. Washington, DC: US Dept of Health and Human Services, Public Health Service, National Institutes of Health, National Institute on Alcohol Abuse and Alcoholism; 1993.
2. Volk RJ, Steinbauer JR, Cantor SB, Holzer CE III. The Alcohol Use Disorders Identification Test (AUDIT) as a screen for at-risk drinking in primary care patients of different racial/ethnic backgrounds. *Addiction*. 1997;92:197-206.
3. Piccinelli M, Tessari E, Bortolomasi M, et al. Efficacy of the alcohol use disorders identification test as a screening tool for hazardous alcohol intake and related disorders in primary care: a validity study. *BMJ*. 1997;314:420-424.
4. Adams WL, Barry KL, Fleming MF. Screening for problem drinking in older primary care patients. *JAMA*. 1996;276:1964-1967.
5. Reid MC, Fiellin DA, O'Connor PG. Hazardous and harmful drinking in primary care. *Arch Intern Med*. 1999;159:1681-1689.
6. *The Physicians' Guide to Helping Patients With Alcohol Problems*. Washington, DC: US Dept of Health and Human Services, Public Health Service, National Institutes of Health, National Institute on Alcohol Abuse and Alcoholism; 1995.
7. Samet JH, Rollnick S, Barnes H. Beyond CAGE: a brief clinical approach after detection of substance abuse. *Arch Intern Med*. 1996;156:2287-2293.
8. Friedmann PD, Saitz R, Samet JH. Management of adults recovering from alcohol or other drug problems. *JAMA*. 1998;279:1227-1231.
9. World Health Organization. *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines*. Geneva, Switzerland: World Health Organization; 1992.
10. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*. Washington, DC: American Psychiatric Press; 1994.
11. First MB, Spitzer RL, Gibbon M, Williams JBW. *Structured Clinical Interview for* DSM-IV *Axis I Disorders*. New York: Biometrics Research Dept, New York State Psychiatric Institute; 1995.
12. Reid MC, Lachs MS, Feinstein AR. Use of methodological standards in diagnostic test research: getting better but still not good. *JAMA*. 1995;274:645-651.
13. Jaeschke R, Guyatt GH, Sackett DL, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature, III: how to use an article about a diagnostic test, B: what are the results and will they help me in caring for my patients? *JAMA*. 1994;271:703-707.
14. Meade MO, Richardson WS. Selecting and appraising studies for a systematic review. *Ann Intern Med*. 1997;127:531-537.
15. Feinstein AR. *Clinical Epidemiology: The Architecture of Clinical Research*. Philadelphia, Pa: WB Saunders Co; 1985.
16. Cyr MG, Wartman SA. The effectiveness of routine screening questions in the detection of alcoholism. *JAMA*. 1988;259:51-54.
17. 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.
18. King M. At risk drinking among general practice attenders: validation of the CAGE questionnaire. *Psychol Med*. 1986;16:213-217.
19. Aithal GP, Thornes H, Dwarakanath AD, Tanner AR. Measurement of carbohydrate-deficient transferrin (CDT) in a general medical clinic: is this test useful in assessing alcohol consumption? *Alcohol Alcohol*. 1998;33:304-309.
20. Sillanaukee P, Aalto M, Seppa K. Carbohydrate-deficient transferrin and conventional alcohol markers as indicators for brief intervention among heavy drinkers in primary health care. *Alcohol Clin Exp Res*. 1998;22:892-896.
21. Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. *Arch Intern Med*. 1998;158:1789-1795.
22. Bradley KA, McDonell MB, Bush K, Kivlahan DR, Diehr P, Fihn SD. The AUDIT alcohol consumption questions: reliability, validity, and responsiveness to change in older male primary care patients. *Alcohol Clin Exp Res*. 1998;22:1842-1849.
23. Bradley KA, Bush KR, McDonell MB, Malone T, Fihn SD. Screening for problem drinking: comparison of CAGE and AUDIT. *J Gen Intern Med*. 1998;13:379-388.
24. Taj N, Devera-Sales A, Vinson DC. Screening for problem drinking: does a single question work? *J Fam Pract*. 1998;46:328-335.
25. 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 persons with harmful alcohol consumption—I. *Addiction*. 1993;88:349-362.
26. Escobar F, Espi F, Canteras M. Diagnostic tests for alcoholism in primary health care: compared efficacy of different instruments. *Drug Alcohol Depend*. 1995; 40:151-158.
27. Gabrynowicz JW, Watts DJ. Early diagnosis of alcoholism: a quick and reliable technique. *Aust Fam Physician*. 1981;10:893-898.
28. Rumpf HJ, Hapke U, Hill A, John U. Development of a screening questionnaire for the general hospital and general practices. *Alcohol Clin Exp Res*. 1997;21: 894-898.
29. Rydon P, Redman S, Sanson-Fisher RW, Reid AL. Detection of alcohol-related problems in general practice. *J Stud Alcohol*. 1992;53:197-202.
30. Barry KL, Fleming MF. The Alcohol Use Disorders Identification Test (AUDIT) and the SMAST-13: predictive validity in a rural primary care sample. *Alcohol Alcohol*. 1993;28:33-42.
31. Isaacson JH, 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;9:550-553.
32. Schmidt A, Barry KL, Fleming MF. Detection of problem drinkers: the Alcohol Use Disorders Identification Test (AUDIT). *South Med J*. 1995;88:52-59.
33. Morton JL, Jones TV, Manganaro MA. Performance of alcoholism screening questionnaires in elderly veterans. *Am J Med*. 1996;101:152-159.
34. Steinbauer JR, Cantor SB, Holzer CE, Volk RJ. Ethnic and sex bias in primary care screening tests for alcohol use disorders. *Ann Intern Med*. 1998;129:353-362.
35. Buchsbaum DG, Buchanan RG, Centor RM, Schnoll SH, Lawton MJ. Screening for alcohol abuse using CAGE scores and likelihood ratios. *Ann Intern Med*. 1991; 115:774-777.
36. 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;40:662-665.
37. Volk RJ, Cantor SB, Steinbauer JR, Cass AR. Item bias in the CAGE screening test for alcohol use disorders. *J Gen Intern Med*. 1997;12:763-769.
38. Chan AW, 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;6:123-135.
39. Brown RL, Rounds LA. Conjoint screening questionnaires for alcohol and other drug abuse: criterion validity in a primary care practice. *WMJ*. 1995;94:135-140.
40. Fleming MF, Barry KL. A three-sample test of a masked alcohol screening questionnaire. *Alcohol Alcohol*. 1991;26:81-91.
41. Cherpitel CJ. Performance of screening instruments for identifying alcohol dependence in the general population, compared with clinical populations. *Alcohol Clin Exp Res*. 1998;22:1399-1404.
42. Fleming MF, Barry KL. The effectiveness of alcoholism screening in an ambulatory care setting. *J Stud Alcohol*. 1991;52:33-36.
43. Cleary PD, Miller M, Bush BT, Warburg MM, Delbanco TL, Aronson MD. Prevalence and recognition of alcohol abuse in a primary care population. *Am J Med*. 1988;85:466-471.
44. Barry KL, Fleming MF. Computerized administration of alcoholism screening tests in a primary care setting. *J Am Board Fam Pract*. 1990;3:93-98.
45. Schorling JB, Willems JP, Klas PT. Identifying problem drinkers: lack of sensitivity of the two-question drinking test. *Am J Med*. 1995;98:232-236.
46. Moran MB, Naughton BJ, Hughes SL. Screening elderly veterans for alcoholism. *J Gen Intern Med*. 1990;5:361-364.
47. Buchsbaum DG, Welsh J, Buchanan RG, Elswick RK Jr. Screening for drinking problems by patient self-report: even "safe" levels may indicate a problem. *Arch Intern Med*. 1995;155:104-108.
48. 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;7:517-521.
49. Broadhead WE, Leon AC, Weissman MM, et al. Development and validation of the SDDS-PC screen for multiple mental disorders in primary care. *Arch Fam Med*. 1995;4:211-219.
50. Leon AC, Olfson M, Weissman MM, et al. Brief screens for mental disorders in primary care. *J Gen Intern Med*. 1996;11:426-430.
51. Hore BD, Alsafar J, Wilkins RH. An attempt at criterion-oriented validation of an alcoholism questionnaire in general practice. *Br J Addict Alcohol Other Drugs*. 1977;72:19-22.
52. Kitchens JM. Does this patient have an alcohol problem? *JAMA*. 1994;272:1782-1787.
53. Reid MC, Tinetti ME, Brown CJ, Concato J. Physician awareness of alcohol use disorders among older patients. *J Gen Intern Med*. 1998;13:729-734.