
Self-Report Screening for Alcohol Problems Among Adults

Gerard J. Connors, Ph.D.,* and Robert J. Volk, Ph.D.†

*Research Institute on Addictions, University at Buffalo, Buffalo, NY

†Baylor College of Medicine, Houston, TX

Alcohol abuse and alcoholism are serious public health problems estimated to affect approximately 7 percent of the U.S. population (Grant et al. 1994), but many individuals with such problems remain undetected. Also undetected are many individuals who do not meet diagnostic criteria for alcohol abuse or alcohol dependence, but who nevertheless are experiencing negative consequences associated with their use of alcohol or are at risk for such consequences (Institute of Medicine 1990). This is unfortunate for several reasons. First, their continued drinking holds significant potential for further alcohol-related negative consequences. Second, it is not possible to refer such drinkers for appropriate services until they are detected. Particularly noteworthy in this regard would be persons experiencing mild to moderate levels of alcohol problems, who respond well to secondary prevention interventions. As such, there is a need to develop and apply techniques to screen for alcohol use disorders. Fortunately, much work has occurred in this area, and this chapter focuses on a variety of issues and measures relevant to the identification of adults with alcohol-related problems. (The topic of screening among adolescents is covered in the chapter by Winters.)

OVERVIEW OF CHAPTER

The first section of this chapter provides a working definition of screening, identifies the goals of

screening, discusses the distinction between screening and assessment, and comments on screening in relation to the treatment process. The next section addresses issues in the evaluation of screening measures, such as sensitivity and specificity. The topic of the validity of self-report data also is addressed. An overview of self-report screening measures is presented, followed by a discussion of guidelines for the selection and use of screening measures, a summary of studies that have compared measures, and some general suggestions regarding screening. The chapter closes with a description of future directions and needs for clinical research in the area of screening.

Definition of Screening

Definitions for the term *screening* are numerous, ranging from the narrowest to broadest breadth of focus or coverage. For purposes of this chapter, the term will be used to represent the skillful use of empirically based procedures for identifying individuals with alcohol-related problems or consequences or those who are at risk for such difficulties.

Empirically based procedures may include biological markers as well as self-report techniques. For example, elevated levels of gamma-glutamyltransferase (GGT) and mean corpuscular volume (MCV) have been used as a screen for excessive alcohol consumption (see Leigh and

Skinner 1988; Rosman and Lieber 1990; and the chapter by Allen et al. in this *Guide* for more detail on such laboratory tests). However, this chapter will focus on self-report screening procedures.

The definition of screening proposed here does not include diagnosis. Screening measures are not intended to provide a diagnosis; assessment for purposes of diagnosis occurs in subsequent stages of evaluation (see the chapter by Maisto et al. in this *Guide* for more detail on diagnostic procedures). The distinction between screening and assessment is discussed below.

Goals of Screening

Having identified a working definition of screening, it makes sense to step back for a moment and specify the goals or objectives of screening. A primary objective is to detect individuals with alcohol problems. In this regard, the population of interest is persons who are not yet addressing their alcohol use disorders. A companion objective is setting the stage for subsequent assessment and, as warranted, interventions. The broader benefit to society is to minimize the human and economic costs of alcohol abuse through detection and intervention, especially early detection so that interventions can be applied as soon as possible.

Distinguishing Between Screening and Assessment

Screening is designed to identify persons experiencing an alcohol use problem. An abnormal or positive screening result may thus “raise suspicion” about the presence of an alcohol use problem, while a normal or negative result should suggest a low probability of an alcohol use problem. Screening measures are not designed (if for no other reason than because of their brevity) to explicate the nature and extent of such problems. By contrast, assessment procedures *are* designed to explore fully the nature and extent of

a person’s problems with alcohol (see the chapter by Maisto et al.). Such assessment information can be used to determine whether the person meets the criteria for a particular diagnostic category, such as alcohol abuse or alcohol dependence, depending on the nomenclature system being applied.

Screening in Relation to the Treatment Process

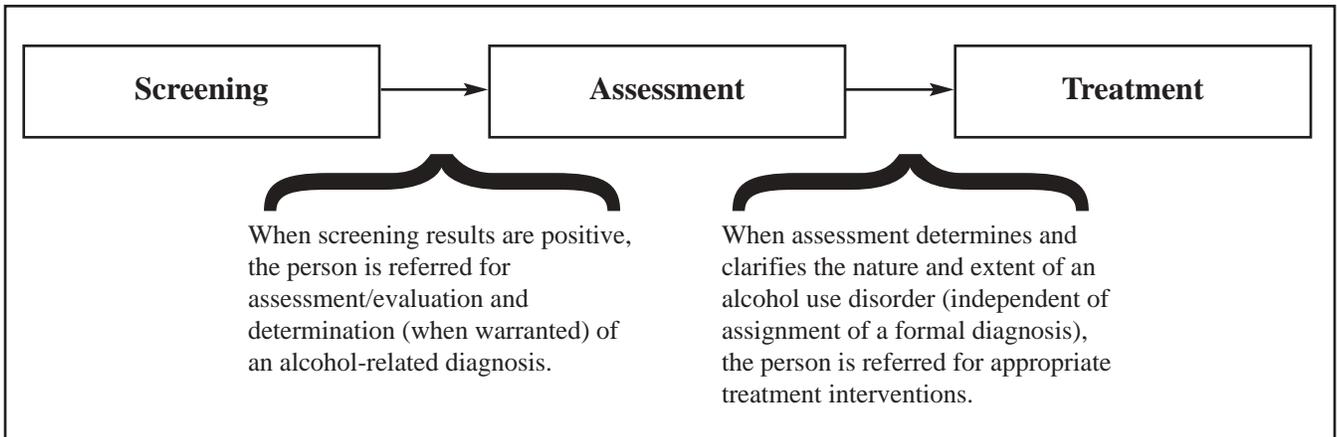
Screening ideally should occur in a manner that facilitates subsequent assessment or referral for assessment among persons identified as positive on the screening measure. For example, screening plans should include sensitive procedures for the communication of screening results in a manner that maximizes the likelihood that the individual will follow through with assessment. Further, any screening system will require procedures for the actual assessment of those identified as positive (through subsequent assessment at the same location or through a referral). The benefits of screening to the individual and society ultimately will be a function of the extent to which identified persons subsequently address their drinking problems. A staging process for these events is depicted in figure 1. Adapted from Allen (1991), the figure shows the connections between screening, assessment, and treatment.

ASSESSING SCREENING MEASURES

Approaches to Evaluating Measures

There are a variety of dimensions along which one can determine the strengths of a particular screening measure. Because of their relevance to evaluating measures and making determinations regarding the utility of specific measures for particular purposes, settings, or populations, it is important to identify and describe these dimensions: sensitivity, specificity, predictive value,

FIGURE 1.—Interrelationships between stages of screening, assessment, and treatment



likelihood ratios, and receiver operating curves. The “gold standard” by which a screening test is evaluated (called the reference test or criterion) generally is a full diagnostic evaluation.

Sensitivity

The sensitivity (or true positive rate) of a test concerns its ability to identify people with the disorder in question, in this case alcohol problems. Stated differently, sensitivity reflects the proportion of persons with alcohol use disorders correctly identified (“true positives”) by the test. Consistent with this definition, a sensitive test is one that provides a minimum of false negatives (i.e., persons with alcohol problems who are not detected by the screening measure).

Table 1 depicts the relationships between test results and alcohol problems. Four outcomes are possible (true positives, false positives, false negatives, and true negatives) for the crossing of the test results (negative or positive) with the disorder (present or absent). Using this grid, sensitivity would be calculated by dividing the true positive cases by the total number of persons with an alcohol use disorder ($a/a + c$). Similarly, the false negative rate, or 1 minus the sensitivity of the test, would be calculated by dividing the false negative cases by the total number of persons with a disorder.

Specificity

The specificity (or true negative rate) of a test refers to its ability to accurately identify people who do not have an alcohol use disorder. As such, specificity reflects the proportion of non-alcohol abusers correctly identified (“true negatives”). Accordingly, a specific test provides a minimum of false positives (i.e., non-alcohol abusers identified by the screening test as alcohol abusers). Referring again to table 1, specificity would be calculated by dividing the true negative cases by the total number of non-alcohol abusers ($d/b + d$). Similarly, the false positive rate, or 1 minus specificity, would be calculated by dividing the false positive cases by the total number of non-alcohol abusers ($b/b + d$).

TABLE 1.—Possible outcomes in screening for alcohol use disorders

Result of screening measure	Alcohol use disorder	
	Present	Absent
Positive	True positives (a)	False positives (b)
Negative	False negatives (c)	True negatives (d)

As a general rule, screening tests tend to emphasize maximizing sensitivity over specificity. This logic is apparent when the purpose of screening is considered. Screening is done on unselected groups (e.g., asymptomatic primary care patients) for the purpose of identifying cases where there is a heightened suspicion of a disorder. For people screening positive, additional testing is done to determine the presence and severity of a problem. The costs of using self-report screening tests are fairly minimal compared with, for example, biochemical tests, and thus specificity becomes less of a concern. Clearly, though, specificity is an important concern as it relates to the resources used to evaluate people who screen positive but do not have an alcohol disorder.

Predictive Value

In general, good screening tests when negative should “rule out” an alcohol use disorder, and when positive should “rule in” a disorder such that assessment is warranted. A useful statistic in evaluating screening tests is called positive predictive value. This refers to the proportion of persons identified as positive on the screening test who actually have the disorder. Clinically, positive predictive value represents the probability of an alcohol use disorder given a positive test result. Referring to table 1, the likelihood that a person with a positive test result actually has an alcohol problem is calculated by dividing the true positives by the number of positives identified by the screening test ($a/a + b$). It should be noted that as the prevalence of the disorder in the population being screened increases, the positive predictive value of the measure increases as well. A related concept is the “false alarm rate,” which is the probability that a person testing positive does not have an alcohol use disorder ($b/a + b$).

Negative predictive value represents the probability that a person does not have an alcohol use

disorder following a negative test result (calculated as $d/c + d$ from table 1). Yet, the more interesting clinical question is, given a negative test result, does this patient still have an alcohol use disorder? The “false reassurance rate,” or 1 minus negative predictive value, represents the probability that a patient has an alcohol use disorder given a negative test result (calculated as $c/c + d$ from table 1). As the prevalence of the disorder in the population goes down, the false reassurance rate also goes down.

Likelihood Ratios

The method of likelihood ratios to describe the accuracy of a screening test has been touted as quicker and more powerful than the sensitivity/specificity strategy. Increasingly, studies of the characteristics of alcohol screening tests are using likelihood ratios as a summary measure. According to Sackett (1992), a likelihood ratio reflects the odds that a positive finding on a screening test would occur in a person with, as opposed to a person without, an alcohol use disorder. He described the significance of different likelihood ratios as follows:

When a finding’s likelihood ratio is above 1.0, the probability of disease goes up (because the finding is *more* likely among patients with, than without, the disorder); when the likelihood ratio is below 1.0, the probability of disease goes down (because the finding is *less* likely among patients with, than without, the disorder); finally, when the likelihood ratio is close to 1.0, the probability of disease is unchanged (because the finding is *equally* likely in patients with, and without, the disorder). (Sackett 1992, pp. 2643–2644, emphasis in original)

The calculation of the likelihood ratio for a positive test result is based on sensitivity and specificity, as follows:

$$\frac{\text{sensitivity}}{1 - \text{specificity}}$$

The likelihood ratio is thus a single number (or ratio) summarizing the characteristics of a test. Proponents of likelihood ratios have argued that they are easily remembered and provide a shorthand method for calculating posttest (posterior) probabilities (Fagan 1975). To do so, it is necessary to reexpress the prior probability as odds using the following formula:

$$\text{Prior Odds} = \text{Probability} / (1 - \text{Probability})$$

For example, a probability of 0.50 is equivalent to an odds of 1.0, interpreted as “one to one” (or 1:1). Thus, for every one patient with the disease there is one patient without the disease (and hence, the probability of disease is 0.50).

Positive predictive value (or posterior probability of a positive result) is calculated by multiplying the prior odds and likelihood ratio and reexpressing the posterior odds as a probability. The following two equations describe these calculations:

$$\text{Posterior Odds} = \text{Prior Odds} \times \text{Likelihood Ratio}$$

$$\text{Posterior Probability} = \text{Posterior Odds} / (1 + \text{Posterior Odds})$$

While likelihood ratios are often used to describe the characteristics of a test, their clinical use has been more limited. One primary limitation of likelihood ratios is the need to reexpress prior and posterior probabilities as odds in calculating predictive value (Dujardin et al. 1994). More information on likelihood ratios and their uses is provided by Feinstein (1985) and Sackett (1992).

Receiver Operating Curves

Receiver operating curves are used to determine optimal cutoff scores for use with a particular screening measure, and in general to describe the overall characteristics of a measure through determining the area under the receiver operating characteristic curve. Changing the test’s cutoff, naturally, has implications for its sensitivity, specificity, and positive predictive value. For example, lowering the cutoff for a screening test generally will identify a greater number of positive test results. Such a strategy typically will result in greater sensitivity, but at the same time it will reduce the test’s specificity. An excellent example of the effect of using different cutoff points for several screening measures (e.g., CAGE, Michigan Alcoholism Screening Test [MAST], T-ACE, and TWEAK) was presented by Russell et al. (1994).

Self-Report Validity and Screening Tests

Although some researchers and clinicians have argued that information from self-reports on alcohol-related variables is suspect (e.g., alcohol abusers will deny they have problems), many others believe these reports can be valid and useful in the screening as well as assessment and treatment of alcohol abusers. This controversy over self-reports has been discussed in greater detail by Babor et al. (1987), Maisto et al. (1990), and Sobell and Sobell (1990).

Clinical researchers in the alcohol field generally accept the idea that the degree of confidence in self-report data increases when information is collected in multiple modes and under circumstances shown to enhance self-reports regarding alcohol use (Babor et al. 1987). For example, the accuracy of self-reports may decrease as a function of recent alcohol consumption, concurrent psychiatric problems, physical and cognitive impairments, the absence of assurances of confidentiality, and an

ambiguous or strained relationship between the person administering the screening measure and the person taking it (see Skinner 1984). Additional considerations relevant to minimizing response bias and maximizing the validity of self-reports include providing clear instructions about the screening task, engaging the person in the process, and ensuring that screening administrators are trained and facile in the task (Babor et al. 1987). Taken together, these and other strategies, depending on the context of the screening endeavor, will yield greater confidence in the self-reports provided by those being screened for alcohol problems.

OVERVIEW OF SCREENING MEASURES

There is no shortage of screening measures available for clinicians and researchers, and a culling of the available measures to a manageable number was performed for purposes of this chapter. Application of the inclusion criteria for this *Guide* (see Allen's "Introduction") yielded a core group of 14 screening measures. Tables 2A and 2B provide descriptive and administrative information on these measures, including examples of groups the measure has been used with, availability of normative data, format, number of items, and time needed to administer the measure. (Table 2A indicates whether norms are available generally as well as for particular subgroups.) Availability of psychometric data, including various types of reliability and validity, is indicated in table 3; see the appendix for more detail.

All of the measures listed in tables 2A and 2B are available for use with adults, and five of them were developed for use with adolescents as well. The measures range in length from very few items (such as the 4-item CAGE) to the 350-item Computerized Lifestyle Assessment (CLA). Six of the screening measures listed in the tables include 10 or fewer items (Alcohol Use Disorders Identification Test [AUDIT], CAGE, Five-Shot Questionnaire, Rapid Alcohol Problems Screen,

T-ACE, and TWEAK). Several of the measures include two or more distinct scales, should such further information be of utility in a particular screening endeavor.

The majority of measures are available for use in a pencil-and-paper self-administered format, but other options are present. Several measures (e.g., AUDIT, CAGE, and MAST) can be used in an interview format, and several measures (e.g., Addiction Potential Scale, AUDIT, CAGE, Drug Use Screening Inventory, Self-Administered Alcoholism Screening Test [SAAST], Substance Abuse Subtle Screening Inventory, and TWEAK) have been adapted for computerized assessment. Regardless of format, most measures can be completed in under 15 minutes, and six can be completed in just 1 or 2 minutes. Scoring of the majority of the measures likewise requires relatively little time.

Overall, the material presented in the tables shows that screening measures have considerable variability in length and potential applicability to particular screening contexts. The process of evaluating and selecting a particular screening measure requires consideration of a number of factors, and these are addressed in the following section.

SELECTION OF MEASURES

It is not possible to make definitive statements on the selection of a screening measure because screening endeavors can vary dramatically along a number of dimensions, such as the population involved, the amount of time available for screening, the setting, and the goals of the screening. However, it is possible to provide guidelines and suggestions. This section provides guidelines for selecting and using a screening measure, summarizes studies that have compared screening measures, and makes some general suggestions regarding screening for alcohol problems. It is important to remember that these guidelines and

TABLE 2A.—Self-report screening measures: Descriptive information

Measure	Target population	Groups used with	Norms avail.?	Normed groups	No. items (no. subscales)
AAS	Adults		Yes	Normals; substance abusers; psychiatric patients	13
APS	Adults		Yes	Normals; alcohol/drug abusers; psychiatric patients	39
AUDIT ¹	Adults	Primary care, ER, surgery, psychiatric patients; DWI offenders; criminals in court, jail, and prison; enlisted men in Armed Forces; workers in EAPs and industrial settings	Yes	Heavy drinkers; alcoholics	10 (3)
CAGE	Adults and adolescents > 16 yrs.	General medical population in a primary care setting	Yes		4
CLA	Adults and adolescents		Yes		350 (20)
DUSI-R	Adults and adolescents > 16 yrs.; youth 10–16 yrs.	Known or suspected alcohol/drug users; matching specific treatments to specific problems	Yes		159 (11)
Five-Shot Questionnaire	Adults	Male early-phase heavy drinkers	Yes	Moderate/heavy drinkers; alcoholics	5
Mac	Adults	Alcoholics likely to deny problems with drinking when asked directly	Yes	Women; alcoholics with collateral drug problems	49

TABLE 2A.—Self-report screening measures: Descriptive information (continued)

Measure	Target population	Groups used with	Norms avail.?	Normed groups	No. items (no. subscales)
MAST ²	Adults and adolescents	Alcoholics, medical patients, psychiatric patients	Yes		25
RAPS4	Adults	ER and primary care settings	No		4
SAAST	Adults	General medical patients	Yes	Gender; age	35 (2)
SASSI	Adults and adolescents	Adolescents (12–18 yrs.); inpatient and outpatient adults	Yes		Adults 93 (10); adolescents 100 (12)
T-ACE	Adults	Pregnant women	Yes	African American inner-city women attending antenatal clinic	4
TWEAK	Adults	Women	Yes	African American gravidas in inner-city clinic; M&F general population; M&F alcoholic patients; M&F outpatients	5

Note: The measures are listed in alphabetical order by full name; see the text for the full names. Information in the table is based primarily on material provided by the developers of the measures; see the appendix for more detail. DWI = driving while intoxicated; EAPs = employee assistance programs; ER = emergency room; M&F = male and female.

¹ Also available is a 3-item version called the AUDIT-C (see Piccinelli et al. 1997 and Gordon et al. 2001).

² Briefer versions of the MAST are available: the 10-item Brief MAST (Pokorny et al. 1972); the 13-item Short MAST (SMAST) (Selzer et al. 1975); and the 9-item modified version of the Brief MAST, called the Malmö modification (Mm-MAST) because it was first used in the city of Malmö (Kristenson and Trell 1982). Also available is a geriatric version of the MAST, called the MAST-G (Mudd et al. 1993). Magruder-Habib et al. (1982) developed a MAST variant called the VAST, designed to distinguish between lifetime and current problems with alcohol.

TABLE 2B.—Self-report screening measures: Administrative information

Measure	Format options ¹	Time to administer (minutes)	Computer scoring avail.?	Fee for use? ²
AAS	P&P SA; computer SA	5	Yes	Yes
APS	P&P SA; computer SA	10	Yes	Yes
AUDIT ³	P&P SA; interview; computer SA	2	No	No
CAGE	P&P SA; interview; computer SA	<1	No	No
CLA	Computer SA	20–30	Yes	Yes
DUSI-R	P&P SA; interview; computer SA	20	Yes	Yes
Five-Shot Questionnaire	P&P SA	1	No	Yes
Mac	P&P SA; computer SA	10	No	Yes
MAST ⁴	P&P SA; interview	8	No	No
RAPS4	Interview	1	No	No
SAAST	P&P SA; computer SA	5	Yes	Unknown
SASSI	P&P SA; computer SA	10–15	Yes	Yes
T-ACE	P&P SA; interview	1	No	Unknown
TWEAK	P&P SA; interview; computer SA	<2	No	No

Note: The measures are listed in alphabetical order by full name; see the text for the full names. Information in the table is based primarily on material provided by the developers of the measures; see the appendix for more detail. P&P = pencil and paper; SA = self-administered.

¹ Most of the self-administered tests can be supervised and scored by office or clinic staff in relatively brief periods of time.

² Information on fees was not always clear, so potential users should confirm whether there are fees before using any of these measures.

³ Also available is a 3-item version called the AUDIT-C (see Piccinelli et al. 1997 and Gordon et al. 2001).

⁴ Briefer versions of the MAST are available: the 10-item Brief MAST (Pokorny et al. 1972); the 13-item Short MAST (SMAST) (Selzer et al. 1975); and the 9-item modified version of the Brief MAST, called the Malmö modification (Mm-MAST) because it was first used in the city of Malmö (Kristenson and Trell 1982). Also available is a geriatric version of the MAST, called the MAST-G (Mudd et al. 1993). Magruder-Habib et al. (1982) developed a MAST variant called the VAST, designed to distinguish between lifetime and current problems with alcohol.

TABLE 3.—Availability of psychometric data on self-report screening measures

Measure	Reliability			Validity		
	Test-Retest	Split-half	Internal consistency	Content	Criterion	Construct
AAS	•		•	•	•	
APS	•		•		•	
AUDIT	•		•	•	•	•
CAGE					•	
CLA	•		•	•	•	•
DUSI-R	•	•	•	•	•	•
Five-Shot Questionnaire					•	
Mac	•		•		•	•
MAST	•		•	•	•	
RAPS4					•	
SAAST			•		•	
SASSI	•				•	
T-ACE					•	
TWEAK					•	

Note: The measures are listed in the same order as in table 2; see the text for the full names of the instruments.

suggestions need to be evaluated carefully in the context of the particular setting and context in which the screening will occur.

Guidelines for Selecting and Using Measures

There are four central questions that need to be addressed in selecting a screening measure:

- The goals of the screening
- The characteristics of the measure for the target population
- The time and resources available for conducting the screening
- The resources available for scoring the screening measure and providing feedback/referral for positive cases

Identifying the goals of screening in a particular situation might appear straightforward. Indeed,

all screening endeavors on some level are designed to detect alcohol problems among those tested. However, the degree of sensitivity and specificity desired will affect the selection of the measure. While one investigator may want to focus on maximizing sensitivity and thus identify as many true positives as possible, another investigator may want to key on specificity and thus maximize the likelihood that persons identified as positive are actually experiencing an alcohol problem.

The characteristics of the screening measure for use with the target population are also an important consideration in selecting a measure. Generally, a measure with high sensitivity is desirable, and ideally this has been demonstrated in screening populations similar to the target group. Measures with high likelihood ratios have the benefit of both high sensitivity and specificity, and may be effective in both ruling in and ruling out

alcohol use problems. Similar information can be gained from the area under the characteristic receiver operating curve, although this estimate is only a global measure of a measure's characteristics, and it is desirable to consider sensitivity and specificity at a given cutoff point.

The amount of time available for performing the screening should not be a major impediment to its conduct. Several screening measures can be completed in just a couple of minutes. For measures that take more time to complete, one must weigh the relative benefits or advantages of the measures against the time factor. The resources required to facilitate screening should also not be a major impediment. The majority of available measures can be administered by clinical or administrative staff with a minimal degree of training (e.g., clerical staff), and many measures can be self-administered. In addition, several measures have been developed for computer administration.

Finally, one must evaluate the resources available for scoring and interpreting the screening data collected and for acting on the results. Conveniently, a host of measures that can be scored and evaluated in just a few minutes are available. Since screening is intended to detect persons with alcohol problems, resources to provide feedback and referral for evaluation and assessment will be needed. The sensitivity versus specificity emphasis of a given measure will have implications for the amount of resources necessary for subsequent feedback and referral of positive cases.

Contrasts Among Screening Measures

Another resource for selecting a screening measure is data on direct comparisons between measures. A number of such efforts, using a variety of screening measures in a range of settings, have been conducted (e.g., Russell et al. 1994; Maisto et al. 1995; Cherpitel 1997; Clements 1998; Seppa et al. 1998; Steinbauer et

al. 1998; Cherpitel and Borges 2000; Aertgeerts et al. 2001). Maisto et al. (1995), for example, reviewed research involving direct contrasts of self-report screening measures for alcohol problems in a variety of settings. Among their conclusions was that the MAST generally was more sensitive than the CAGE, although the CAGE may perform better than the MAST with elderly primary care patients, and that the CAGE and the Short MAST performed comparably. They noted that the CAGE is particularly popular in primary care settings.

Cherpitel (1997) described the relative strengths of the AUDIT, the TWEAK, the CAGE, and the Brief MAST in population subgroups. Among the conclusions were that the AUDIT and the TWEAK showed greater sensitivity than the CAGE or the Brief MAST and that the instruments were more sensitive for men than for women. However, notable subgroup patterns emerged. The AUDIT and the TWEAK were equally sensitive among African Americans, while the TWEAK was more sensitive than the AUDIT among Whites. Further, the sensitivity of the AUDIT and the TWEAK among African Americans and White men did not differ, while among women, the AUDIT was more sensitive among African Americans and the TWEAK more sensitive among Whites.

Steinbauer et al. (1998) administered the CAGE, the SAAST, and the AUDIT to patients at an adult family medicine clinic. They were particularly interested in identifying ethnic and/or gender biases in the measures. They found that the CAGE and the SAAST showed poorer performance than the AUDIT in identifying alcohol use disorders among African American men, White women, and Mexican American patients. Each measure showed good discriminability for African American women. Steinbauer et al. concluded by recommending that the AUDIT be used in primary health care settings, including those serving multi-ethnic communities. In another report comparing

measures (including the AUDIT, the CAGE, and the MAST), Clements (1998) found the AUDIT to be superior at identifying current alcohol dependence among undergraduate students.

The conclusions provided by these reports comparing screening measures may be useful in deliberations involving the choice of specific scales, particularly in terms of matching screening measures according to gender and ethnicity. However, these studies have included only a subset of the measures listed in tables 2A and 2B. Thus, their findings should not necessarily be used to choose any of the measures they surveyed over the remainder of measures listed in the tables.

Investigations also have been conducted on the use of screening measures (including several of those described in tables 2A and 2B) composed of items selected from other scales and on the use of screens including only one or two questions. The four-item T-ACE, for example, includes three items from the CAGE along with an item on tolerance, and the five-item TWEAK includes three T-ACE items and two MAST items. As another example, Cherpitel (1995) developed the Rapid Alcohol Problems Screen for use in emergency room settings. This five-item measure is composed of two questions from the TWEAK, two from the AUDIT, and one from the Brief MAST. A four-item version, called the RAPS4, has also been developed (Cherpitel 2000). Seppa and colleagues (1998) developed the Five-Shot Questionnaire, which includes two items from the AUDIT and three from the CAGE. In evaluating the questionnaire with middle-aged men attending a health screening, Seppa et al. found the Five-Shot Questionnaire to be efficient in differentiating between moderate and heavy drinkers. In an even briefer approach, Cyr and Wartman (1988) recommended two screening questions (“Have you ever had a drinking problem?” and “When was your last drink?”); Taj et al. (1998) proposed the use of a single question (“On any single occasion during the past 3 months, have you had more than 5 drinks containing alcohol?”). Williams and Vinson

(2001) also proposed a single question (“When was the last time you had more than X drinks in 1 day?” where X = 4 for women and 5 for men). Brown and colleagues (2001), in an effort to assess both alcohol and other substance abuse, have developed a two-item conjoint screen (TICS). The items are “In the past year, have you ever drunk or used drugs more than you meant to?” and “Have you felt you wanted or needed to cut down on your drinking or drug use in the last year?”

Suggestions

Although, as has been emphasized throughout this chapter, it is important to consider the specific goals, setting, and other factors in selecting a screening measure, there are some general suggestions that can be made regarding screening for alcohol problems. These suggestions (see also Allen et al. 1995 and Maisto et al. 1995) have particular relevance to primary health care settings, where screening for alcohol problems is becoming more frequent.

First, there is a wide array of screening measures that can be recommended generally for use with adults. Although the choice will be dictated, of course, by the specific needs of the program, the AUDIT can be recommended for a variety of settings. It has been shown to possess a number of strengths and advantages. For settings in which a briefer approach is needed, there are several screens available that involve administration of only one or two questions.

Second, screening projects should consider the concomitant use of laboratory tests where available, particularly in health care settings where such tests are routinely performed. Positive results on biochemical tests (e.g., GGT or MCV) may enhance the credibility of self-report screening results when presented to clients. There is some evidence that biochemical markers such as carbohydrate-deficient transferrin (CDT) identify a different spectrum of alcohol use problems than self-report screening tests such as the AUDIT (Hermansson et al. 2000).

Finally, any screening endeavor requires responsive procedures regarding feedback to individuals screened and the making of appropriate referrals for further evaluation and assessment. The establishment of such procedures is a necessary component of the screening process that needs to be in place prior to the actual screening of individuals.

FUTURE DIRECTIONS AND NEEDS

Many screening measures have been developed for use in clinical settings, including primary health care settings. There have been some interesting historical trends in this research, which should be considered as future studies are planned. First, many screening tests share common roots with the CAGE questions and the MAST. There is a fairly extensive literature on the performance of these measures. A second trend has been to develop ever briefer measures, with several single-item measures now being touted. Whether these briefer measures will lead to increased screening, allow for feedback to patients, and provide for optimal management of patients with alcohol use problems has yet to be determined. A final trend has been to emphasize consumption indicators either alone or in combination with other consequence-based or dependence indicators.

Although these advances in screening measures are important, implementation appears to be lagging behind the development and evaluation of measures. Thus, more attention should be paid to strategies and approaches for increasing the use of screening measures in a variety of settings.

There are a number of important research directions that should be considered in enhancing screening for alcohol use problems in clinical settings. Research to date has largely evaluated screening measures in highly protocol-driven, investigator-controlled studies. Research staff are often used to administer the measures, the scoring is provided through the study, and the criterion measure against which the measure is evaluated is also administered by the staff. Such studies might be seen as assessing “efficacy,” or examining the

performance of measures in ideal settings. However, we know comparatively little about how screening measures should be used in real-world clinical settings. Studies are needed to assess the “effectiveness” of screening for alcohol use problems, exploring such factors as the timing of screening, who should administer the screen, who should interpret the results for the clinician and patient, and how the results are to be incorporated with further assessment and management.

A related research concern has to do with the problem of integrating screening within other preventive health care services. For example, in the primary care setting, a routine health examination can include screening for many medical problems and health risk behaviors (e.g., various cancers, hypertension, lipid disorders, seat belt use, bicycle helmet use). Most studies on screening measures have considered a specific measure as part of the instrumentation in a research project rather than integrated within various screening tools administered as part of a routine health maintenance visit. Daepfen et al. (2000) demonstrated that the AUDIT performs well when embedded within a broader general health risk questionnaire. Research is needed to better understand how screening for alcohol use problems can become part of routine health examinations, and how screening tools might be integrated with other health risk assessments. Clearly, it is not enough to argue that screening tests should simply be added as part of the routine office visit without considering competing clinical and administrative demands put upon providers.

Research is also needed on the use of screening measures with specific populations. For example, the Research Institute on Addictions Self Inventory (RIASI) (Nochajski and Wiczorek 1998; Nochajski et al. unpublished manuscript) is a screening measure designed to briefly but accurately determine which driving under the influence (DUI) offenders need to be referred for diagnostic evaluation. The measure, which can be completed and scored in 15 minutes, is being used

to identify DUI arrestees with alcohol and/or other drug problems. The RIASI represents a careful and empirical development of a screening device for use with a particular population. Developed specifically for the New York State Drinking Driver Programs, it is now being used in several State programs for DUI offenders.

A final area for further investigation involves development of testing systems, where combinations of self-report measures, and potentially biochemical markers, are used. Again, research on screening measures has largely considered the performance of measures in isolation or in comparison with other measures. Testing algorithms might be developed where the results of one measure suggest further testing to enhance predictive value and guide assessment.

ACKNOWLEDGMENTS

Preparation of this manuscript was supported in part by grants R01AA11728 and N01AA81015 from the National Institute on Alcohol Abuse and Alcoholism.

REFERENCES

- Aertgeerts, B.; Buntinx, F.; Ansoms, S.; and Fevery, J. Screening properties of questionnaires and laboratory tests for the detection of alcohol abuse or dependence in a general practice population. *Br J Gen Pract* 51:206–217, 2001.
- Allen, J.P. The interrelationship of alcoholism assessment and treatment. *Alcohol Health Res World* 15:178–185, 1991.
- Allen, J.P.; Maisto, S.A.; and Connors, G.J. Self-report screening tests for alcohol problems in primary care. *Arch Intern Med* 155:1726–1730, 1995.
- Babor, T.F.; Stephens, R.S.; and Marlatt, G.A. Verbal report methods in clinical research on alcoholism: Response bias and its minimization. *J Stud Alcohol* 48:410–424, 1987.
- Brown, R.L.; Leonard, T.; Saunders, L.A.; and Pappasoulotis, O. A two-item conjoint screen for alcohol and other drug problems. *J Am Board Fam Pract* 14:95–106, 2001.
- Cherpitel, C.J. Screening for problems in the emergency room: A rapid alcohol problems screen. *Drug Alcohol Depend* 40:133–137, 1995.
- Cherpitel, C.J. Brief screening instruments for alcoholism. *Alcohol Health Res World* 21:348–351, 1997.
- Cherpitel, C.J. A brief screening instrument for problem drinking in the emergency room: The RAPS4. Rapid Alcohol Problems Screen. *J Stud Alcohol* 61:447–449, 2000.
- Cherpitel, C.J., and Borges, G. Screening instruments for alcohol problems: A comparison of cut points between Mexican American and Mexican patients in the emergency room. *Subst Use Misuse* 35:1419–1430, 2000.
- Clements, R. A critical evaluation of several alcohol screening instruments using the CIDI-SAM as a criterion measure. *Alcohol Clin Exp Res* 22:985–993, 1998.
- Cyr, M., and Wartman, S. The effectiveness of routine screening questions in the detection of alcoholism. *JAMA* 259:51–54, 1988.
- Daepfen, J.-B.; Yersin, B.; Landry, U.; Pecoud, A.; and Decrey, H. Reliability and validity of the Alcohol Use Disorders Identification Test (AUDIT) imbedded within a general health risk screening questionnaire: Results of a survey in 332 primary care patients. *Alcohol Clin Exp Res* 24:659–665, 2000.
- Dujardin, B.; Van den Ende, J.; Van Gompel, A.V.; Unger, J.; and Van der Stuyft, P. Likelihood ratios: A real improvement for clinical decision making? *Eur J Epidemiol* 10:29–36, 1994.
- Fagan, T.J. Nomogram for Bayes's Theorem. *N Engl J Med* 293:257, 1975.
- Feinstein, A.R. *Clinical Epidemiology: The Architecture of Clinical Research*. Philadelphia: W.B. Saunders, 1985.
- Gordon, A.J.; Maisto, S.A.; McNeil, M.; Kraemer, K.L.; Conigliaro, R.L.; Kelley, M.E., and Conigliaro, J. Three questions can detect hazardous drinkers. *J Fam Pract* 50:313–320, 2001.
- Grant, B.F.; Harford, T.C.; Dawson, D.A.; Chou, P.; DuFour, M.; and Pickering, R. Prevalence of DSM-IV alcohol abuse and dependence: United States, 1992. *Alcohol Health Res World* 18:243–248, 1994.

- Hermansson, U.; Helander, A.; Huss, A.; Brandt, L.; and Ronnberg, S. The alcohol use disorders identification test (AUDIT) and carbohydrate-deficient transferrin (CDT) in a routine workplace health examination. *Alcohol Clin Exp Res* 24:180–187, 2000.
- Institute of Medicine. *Broadening the Base of Treatment for Alcohol Problems*. Washington, DC: National Academy Press, 1990.
- Kristenson, H., and Trell, E. Indicators of alcohol consumption: Comparisons between a questionnaire (Mm-MAST), interviews and serum gamma-glutamyl transferase (GGT) in a health survey of middle-aged males. *Br J Addict* 77:297–304, 1982.
- Leigh, G., and Skinner, H.A. Physiological assessment. In: Donovan, D.M., and Marlatt, G.A., eds. *Assessment of Addictive Behaviors*. New York: Guilford Press, 1988. pp. 112–136.
- Magruder-Habib, K.; Harris, K.E.; and Fraker, G.G. Validation of the Veterans Alcoholism Screening Test. *J Stud Alcohol* 43:910–926, 1982.
- Maisto, S.A.; McKay, J.R.; and Connors, G.J. Self-report issues in substance abuse: State of the art and future directions. *Behav Assess* 12:117–134, 1990.
- Maisto, S.A.; Connors, G.J.; and Allen, J.P. Contrasting self-report screens for alcohol problems: A review. *Alcohol Clin Exp Res* 19:1510–1516, 1995.
- Mudd, S.A.; Blow, F.C.; Hill, E.M.; Demodananberg, L.; Young, J.P.; and Iacob, A. Differences in symptom reporting between older problem drinkers with and without a history of major depression. *Alcohol Clin Exp Res* 17:489, 1993.
- Nochajski, T.H., and Wieczorek, W.F. Identifying potential drinking-driving recidivists: Do non-obvious indicators help? *Journal of Prevention and Intervention in the Community* 1:69–83, 1998.
- Piccinelli, M.; Tessari, E.; Bortolomasi, M.; Piasere, O.; Semenzin, M.; Garzotto, N.; and Tansella, M. 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* 314:420–424, 1997.
- Pokorny, A.D.; Miller, B.A.; and Kaplan, H.B. The brief MAST: A shortened version of the Michigan Alcoholism Screening Test (MAST). *Am J Psychiatry* 129:342–345, 1972.
- Rosman, A.S., and Lieber, C.S. Biochemical markers of alcohol consumption. *Alcohol Health Res World* 14:210–218, 1990.
- Russell, M.; Martier, S.S.; Sokol, R.J.; Mudar, P.; Bottoms, S.; Jacobson, S.; and Jacobson, J. Screening for pregnancy risk-drinking. *Alcohol Clin Exp Res* 18:1156–1161, 1994.
- Sackett, D.L. The rational clinical examination. A primer on the precision and accuracy of the clinical examination. *JAMA* 267:2638–2644, 1992.
- Selzer, M.L.; Vinokur, A.; and van Rooijen, L. A self-administered Short Michigan Alcoholism Screening Test (SMAST). *J Stud Alcohol* 36:117–126, 1975.
- Seppa, K.; Lepisto, J.; and Sillanaukee, P. Five-shot questionnaire on heavy drinking. *Alcohol Clin Exp Res* 22:1788–1791, 1998.
- Skinner, H.A. Assessing alcohol use by patients in treatment. In: Smart, R.G.; Cappell, H.D.; Glaser, F.B.; Israel, Y.; Kalant, H.; Popham, R.E.; Schmidt, W.; and Sellers, E.M., eds. *Research Advances in Alcohol and Drug Problems*. Vol. 8. New York: Plenum Press, 1984. pp. 183–207.
- Sobell, L.C., and Sobell, M.B. Self-report issues in alcohol abuse: State of the art and future directions. *Behav Assess* 12:77–90, 1990.
- Steinbauer, J.R.; Cantor, S.B.; Holzer, C.E.; and Volk, R.J. Ethnic and sex bias in primary care screening tests for alcohol use disorders. *Ann Intern Med* 129:353–362, 1998.
- Taj, N.; Devera-Sales, A.; and Vinson, D.C. Screening for problem drinking: Does a single question work? *J Fam Pract* 46:328–335, 1998.
- Williams, R., and Vinson, D.C. Validation of a single screening question for problem drinking. *J Fam Pract* 50:307–312, 2001.

