A Review of Research on the Alcohol Use Disorders Identification Test (AUDIT)

John P. Allen, Raye Z. Litten, Joanne B. Fertig, and Thomas Babor

Research on the core version of the Alcohol Use Disorders Identification Test (AUDIT) is reviewed. Sensitivities and specificities of the AUDIT for criteria of current hazardous use and, to a slightly lesser extent, lifetime alcohol dependence are high. In general, AUDIT scores are at least moderately related to other self-report alcohol screening tests. Several studies also show them as correlated with biochemical measures of drinking. Results of the AUDIT have also been associated with more distal indicators of problematic drinking. Indices of internal consistency, including Cronbach’s α and item-total correlations, are generally in the 0.80’s. Future directions for research on the AUDIT are suggested.

Key Words: Alcohol Screening, Alcohol Use Disorders Identification Test, Hazardous Drinking, Alcohol Diagnosis.

Despite its rather recent development, the Alcohol Use Disorders Identification Test (AUDIT)1 has already stimulated an extensive body of research. The AUDIT now ranks as fourth among self-report alcoholism screening measures in terms of research generated. Only the CAGE Questionnaire,2 the Michigan Alcoholism Screening Test (MAST),3 and the MacAndrew Scale4—all of which have been available for much longer periods of time—have attracted more investigation.

In this study, we review the rationale and history of the AUDIT and summarize research findings dealing with its psychometric properties. We conclude our synopsis by identifying a series of questions to which future research efforts could be profitably directed. We elected to conduct such an overview because several important studies on the AUDIT have recently been performed, and have yet to be summarized and integrated with early stage AUDIT projects. Second, the brevity, item coverage, and accuracy of the AUDIT would seem to make the test attractive to primary care health care providers and researchers. We want to afford them objective and comprehensive information on the AUDIT to assist in making decisions on whether or not to incorporate the AUDIT into their professional activities.

HISTORY OF THE MEASURE

The AUDIT was developed in a World Health Organization (WHO) collaborative project performed in primary health care facilities in Australia, Bulgaria, Kenya, Mexico, Norway, and the United States. Nearly 2000 patients were included in the sample. Sixty-four percent were classified as current drinkers, a fourth of whom were further subcategorized as alcoholics. Participants were given a physical examination, including standard blood markers of alcoholism, as well as an extensive questionnaire assessing demographic characteristics, medical history and complaints, use of alcohol and drugs, psychological reactions to alcoholism, as well as an extensive questionnaire assessing demographic characteristics, medical history and complaints, use of alcohol and drugs, psychological reactions to alcohol, problems associated with drinking, family history of alcohol problems, and personal acknowledgment of difficulties related to drinking. Items were selected for the AUDIT from this pool of questions primarily on the basis of correlations with daily alcohol intake, frequency of consuming six or more drinks per drinking episode, and their ability to discriminate hazardous and harmful drinkers. (Hazardous drinkers are persons whose pattern of drinking poses a high risk of future damage to physical or mental health. Harmful drinkers have a pattern of alcohol use that is already resulting in problems.) Items were also chosen according to considerations of face validity, relationships among alcohol-related domains, anticipated usefulness of elicited responses as a focus of therapy, and adequacy of coverage of relevant conceptual domains (i.e., intake, dependence, and adverse consequences of drinking). Finally, attention in item selection was also given to gender appropriateness and cross national generalizability. Questions comprising the AUDIT are presented in the Appendix.

With the exception of the last two items, AUDIT questions allude to the previous year, and responses are weighted between 0 and 4, generally based on frequency of occurrence. The last two items inquire about alcohol-related problems and accord a higher weight for occurrence in the past year and a lower weight for occurrence ever. Sensitivities and specificities of the test were computed for multiple criteria (i.e., heavy average daily consumption, recurrent intoxication, presence of at least one symptom of the alcohol dependence syndrome, diagnosis of alcohol abuse or dependence, and self-perception of a drinking problem). Various potential cut-points in total scores were considered to identify the cut-off value with optimal sensitivity (percentage of positive cases that the test identified) and specificity (percentage of negative cases that the test
identified) to distinguish hazardous and harmful alcohol use. In addition, validity against a criterion of alcoholism was also computed. In the test developmental samples, at a cut-off value of 8 points, sensitivities for the AUDIT for various indices of problematic drinking were generally in the mid 0.90’s and never below the high 0.80’s. Specificities across countries and across criteria averaged in the 0.80’s.¹ Research procedures and results of analyses for the U.S. sample are discussed in detail by Bohn et al.⁵

An instructional handbook for use of the AUDIT in primary care facilities is available.⁶ In addition to presenting a cogent rationale for alcohol screening, the manual summarizes how the measure was derived and offers recommendations for its administration. The AUDIT may be given orally, in writing, or via computer and may be included in a general health risk appraisal assessment. Time for computer-based administration averages ~2 min.⁷ There is no copyright fee for its use.

The AUDIT is unique among self-report screening tests, in that it was constructed based on responses to items from a large multinational sample, used an explicit conceptual-statistical rationale for item selection, emphasizes identification of hazardous drinking rather than long-term dependence and adverse drinking consequences, and focuses primarily on symptoms occurring during the recent past rather than “ever”. In test construction methodology, it may be fairly said to be the most psychometrically sophisticated alcohol screen among at least 22 other candidate self-report instruments.⁸ Nevertheless, the initial validation study had limitations that prompted the developers to encourage further validation research.

SENSITIVITY AND SPECIFICITY OF THE AUDIT AGAINST DIAGNOSTIC CRITERIA

Table 1 summarizes sensitivities and specificities reported for the AUDIT in studies conducted outside the context of the WHO test construction project. At the usual cut-off of 8, most of the projects revealed quite favorable sensitivity and usually lower, but still acceptable, specificity. (In clinical screening for alcohol problems, sensitivity is generally viewed as a more important consideration than specificity, because false-positive screening cases can be subsequently distinguished by more extensive diagnostic procedures.) The study by Isaacson and colleagues⁹ is of particular interest, because it suggests that, as intended, the AUDIT is sensitive to recognizing current problems as opposed to problems with alcohol that occurred only in the past. In this project, none of the subjects with previous, but equally appropriate for males and females. Although Cher- pitel¹⁰ found that, among emergency room patients, the test was somewhat more sensitive and less specific for males than for females, Russell et al.¹² did not observe a gender effect in general hospital patients. Furthermore, area under the receiver operating curve (a geometric representation of the hypothetical validity of the test considering the relative sensitivities and specificities associated with all possible cut-off scores) analyses did not differ as a function of gender for either harmful use or dependence in the study of adult family practice patients.¹³ Figures for sensitivity and specificity of the AUDIT for both genders seem quite favorable.

Only two studies have considered the possible role of ethnicity on test accuracy. Validity of the AUDIT did not differ between Black and White patients in the Cherpetel project.¹⁰ Similarly, areas under the receiver operating curve for harmful and “at-risk” drinking did not differ as a function of ethnicity of the patient as White, Black, or Mexican American.¹³

To date, effects of age have not been systematically analyzed as possible influences on AUDIT sensitivity and specificity. Nevertheless, the reported sensitivity of 57% in Australian inpatients over age 65¹⁴ seems somewhat low and, perhaps, suggestive that the test may be less accurate with older adults. At the other end of the age spectrum, the study with college students suggests that the AUDIT may continue to be quite accurate for young adults if the usual cut-off of 8 is used.¹⁵ All other studies of sensitivity and specificity of the AUDIT have been restricted to largely mid-age range adults. Finally, concurrent drug dependence does not seem to effect the validity of the AUDIT adversely in distinguishing hazardous from nonhazardous drinking.¹⁶

RELATIONSHIP OF THE AUDIT TO OTHER INDICES OF DRINKING

Several studies to establish construct validity of the AU- DIT have been performed. The majority of these have considered issues of convergent validity, concordance with measures that should also be related to the construct measured by the scale. Most prominently, these include efforts to relate AUDIT scores to other self-report and biochemical measures of excessive drinking.

Only the WHO test developmental study¹ seems to have evaluated the relationship of AUDIT scores to commonly used biochemical measures of excessive alcohol use. In the U.S. sample of patients who had been drinking in the past year, the AUDIT correlated in the low-to-moderate range with γ-glutamyltranspeptidase (GGT), aspartate aminotransferase, alanine aminotransferase, and macrocytic volume. Interestingly, the associations tended to be somewhat higher for women than for men.⁵ Of some relevance as well, in a very large scale trial¹⁷ among general hospital patients in Mexico, 44% of those with AUDIT scores of at least 8 also produced serum GGT levels >50 units/liter, the conventional cut-point for GGT. Data on GGT levels for AUDIT negative patients were not given.

Relationships of the AUDIT with alternative self-report measures of drinking tend to be considerably higher than
Table 1. Sensitivities and Specificities of the AUDIT in Unselected Samples against Diagnostic Criteria

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Subjects</th>
<th>AUDIT cut-off</th>
<th>Criterion</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>Australian inpatients aged 65+</td>
<td>8</td>
<td>Hazardous/harmful drinking as defined by ICD and National Health and Medical Research Criteria</td>
<td>57</td>
<td>100</td>
</tr>
<tr>
<td>12</td>
<td>U.S. hospital inpatients aged 18+</td>
<td>8</td>
<td>DSM-III-R diagnosis of alcohol abuse or dependence in the past 12 months</td>
<td>94 (F=100)</td>
<td>80 (F=77)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Alcohol-related ICD-9 discharge diagnosis</td>
<td>92 (F=100)</td>
<td>64 (F=75)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(M=92)</td>
<td>(M=78)</td>
</tr>
<tr>
<td>15</td>
<td>U.S. undergraduate students</td>
<td>7</td>
<td>DSM-III-diagnostic criteria of lifetime alcohol misuse</td>
<td>94</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>94</td>
<td>66 (7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>92</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>88</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>84</td>
<td>71</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>77</td>
<td>69 (7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>70</td>
<td>78</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>61</td>
<td>86</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>49</td>
<td>91</td>
</tr>
<tr>
<td>27</td>
<td>U.S. primary care patients aged 18+</td>
<td>2</td>
<td>DIS-III R criteria for abuse or dependence</td>
<td>100</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>77</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>61</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>55</td>
<td>91</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>50</td>
<td>93</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>38</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>16</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>16</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11</td>
<td>100</td>
</tr>
<tr>
<td>25</td>
<td>Unemployed Norwegians aged 17–63</td>
<td>11</td>
<td>DSM-III-R diagnostic criteria for alcohol misuse or dependence</td>
<td>94</td>
<td>85</td>
</tr>
<tr>
<td>9</td>
<td>U.S. inner city general medical clinic patients aged 18–84</td>
<td>8</td>
<td>DSM-III-R diagnostic criteria for current alcohol abuse or dependence</td>
<td>96</td>
<td>96</td>
</tr>
<tr>
<td>20</td>
<td>U.S. family practice clinic patients aged 19+</td>
<td>2</td>
<td>DSM-III criteria for lifetime and (current) diagnoses</td>
<td>76 (82)</td>
<td>25 (25)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>74 (74)</td>
<td>38 (51)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>72 (68)</td>
<td>51 (51)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>64 (57)</td>
<td>62 (82)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>56 (51)</td>
<td>71 (71)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>53 (47)</td>
<td>85 (83)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>46 (45)</td>
<td>90 (80)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>44 (43)</td>
<td>93 (93)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>35 (41)</td>
<td>96 (95)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>30 (38)</td>
<td>96 (96)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>21 (36)</td>
<td>97 (97)</td>
</tr>
<tr>
<td>10–11</td>
<td>U.S. emergency room patients aged 18+ years who had 12 months</td>
<td>8*</td>
<td>ICD-10 diagnostic criteria for harmful drinking in past 12 months</td>
<td>85 (F=72)</td>
<td>88 (F=97)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(m=93)</td>
<td>(M=77)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(Black = 88)</td>
<td>(Black = 89)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(White = 77)</td>
<td>(White = 84)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(F = 91)</td>
<td>(F = 97)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(M = 90)</td>
<td>(M = 80)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(Black = 85)</td>
<td>(Black = 90)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(White = 76)</td>
<td>(White = 88)</td>
</tr>
</tbody>
</table>

ICD, International Classification of Diseases; DSM, Diagnostic and Statistical Manual; F, female; M, male; DIS, Diagnostic Interview Schedule.
* Original references provide sensitivities and specificities for cut-off scores of 7 and 9 also.

those for biochemical measures. Bohn et al. noted a striking relationship between the AUDIT and the MAST (r = 0.88) for both males and females, and correlations of 0.47 and 0.46 for males and females, respectively, on the MacAndrew Scale, a covert content alcoholism screening test. A high correlation coefficient, 0.78, was also found by Rigmaiden et al. between the AUDIT and the CAGE in ambulatory care patients. Furthermore, these researchers discovered that 88% of ambulatory care patients who scored positive on the CAGE were also identified as positive by the AUDIT, using a cut-off of 7. Even if the AUDIT cut-off were raised to 11, 69% of the CAGE positives would still have been identified by the AUDIT. Hays et al. found moderate Pearson correlations, usually in the 0.60s, for the AUDIT with the Short MAST, CAGE, and a three-item specially constructed “Jellinek” in drink-
ing drivers. AUDIT and MAST scores correlated at 0.77 in a sample of drug abusers in treatment.26

On the other hand, discouraging results have also been reported. Most notably, the correlation between total AUDIT scores and Short MAST scores was found to be only 0.25 among rural primary care patients.20 Furthermore, in a study of patients in Veterans Administration medical and urgent care clinics21 at a cut-off of 8 on the AUDIT, only 51% of the positives also scored positive on the Brief MAST and only 41% of Brief MAST positives were AUDIT positives. Similar discrepancies between positive identifications from the AUDIT and positive cases on both the CAGE and Malmo modification of the MAST22 have been reported by Luckie et al.21 (Comparable data for negative calls were unfortunately not reported by Rigmaiden et al., although agreement was quite high for negative calls in the Luckie et al. study.)

The implications of the Rigmaiden and Luckie projects are particularly important, because both studies reported associations between tests using dichotomized rather than continuous scores on the AUDIT and other screening test. Whereas correlations between alternative screening measures are of some interest, the fundamental concern is the extent to which the tests agree in their final determination of presence or absence of a problem.

Guevara-Arnal et al.17 reported 83% correspondence of AUDIT results as positive or negative between patients’ reports and those from next of kin. The authors unfortunately did not note if the percentages of agreement differed as a function of whether the patient’s AUDIT was positive or negative. Over half of the AUDIT-positive patients also exhibited physical signs of alcohol abuse.

Hays et al.,19 in their study of intoxicated drivers, found additional support for the construct validity of the AUDIT. First offenders scored lower than repeat offenders. Further, self-reported blood alcohol levels at the time of arrest were somewhat correlated with AUDIT scores. Construct validity of the AUDIT is also argued by Seppa et al.23 who reported a correlation of 0.70 between AUDIT scores and self-reported weekly alcohol consumption levels among Finnish research employees. Even among adult male hospital inpatients not diagnosed as alcohol abusers or dependent, the AUDIT demonstrated a sensitivity of 74% and a specificity of 89% to detect patients consuming an average of four or more drinks per day or consumption of at least five drinks per drinking occasion.24

Finally, AUDIT scores have been related to a variety of drinking consequences, attitudes toward drinking, vulnerability to alcoholism, negative affect after drinking, and reasons for drinking. Curiously, relationships between AUDIT scores and such distal measures tend to be somewhat higher for men than for women.3 Whereas the reasons for this are unclear, it may reflect different ages of onset of problem drinking as well as varying patterns and levels of alcohol consumption.

A single study7 seems to have looked at issues of discriminant validity of the AUDIT. In this project, scores on a scale measuring social desirability were slightly inversely related (r's = -0.21 to -0.25) to scores on the AUDIT, as well scores on the Short MAST and the CAGE among drinking drivers. Granted the transparent nature of the content of these self-report screening measures and the likely reluctance of this population to admit alcohol problems, such a finding is reassuring and suggests that individuals completing the test under the more typical circumstances of seeking health care or participating in an anonymous survey for the most part respond to AUDIT questions in a candid manner.

**RELATIONSHIP OF THE AUDIT TO DISTAL CRITERIA**

Not surprisingly, most research to establish validity of the AUDIT has considered the relationship of the total score to current alcohol use/abuse, adverse consequences ascribed directly to drinking, or other specific indicators of problematic drinking. At least two studies, however, have considered possible associations of AUDIT scores with indicators of more global life functioning.

In a particularly innovative project,25 the likelihood of remaining unemployed for at least an additional 2 years was found to be 1.6 times higher for individuals with scores of 8 or more on the AUDIT than for comparable currently unemployed persons with lower scores. As an interesting secondary issue, the researchers also observed that re-employment was accompanied by some decrease in AUDIT scores.

In another prospective study,26 AUDIT scores of ambulatory care patients were found to predict occurrence in the next 2 to 3 years of a physical disorder associated with alcohol use, as well as occurrence of at least one social problem related to drinking. Test scores also predicted health care utilization and future risk of engaging in hazardous drinking (i.e., average daily drinking of >40 g for males and 20 g for females). At a cut-off of 8, the future probability of a social problem was almost seven times and for a medical disorder almost two times greater for positive AUDIT patients than for those with negative scores. These relationships generally held across age and sex. So too, patients with positive AUDIT scores were, respectively, 6.6 and 6.9 times more prone to engage in hazardous drinking and drinking at a level assumed to result in intoxication at least twice a month.

**RELIABILITY OF THE AUDIT**

A small number of studies have reported indices of internal consistency for the AUDIT. Item-total correlations have clearly been demonstrated as highest for the first three AUDIT questions and have generally ranged in the neighborhood of 0.80 among college students,15 rural, general medical practice patients,20 and primary care pa-
tients.²⁷ These items deal with quantity and frequency of drinking. It is unclear if the advantage that these questions evidence is primarily a function of their inherent relationship to the criterion, because the operational definition of hazardous drinking includes level of alcohol consumption or reflects the much higher variances of the first three questions in contrast to the remaining AUDIT items. Unfortunately, none of the cited studies computing item-total correlations seems to have measured the relationship of AUDIT items to the total score without the predictive item included in this sum. Granted that the AUDIT consists of only 10 items, this more conservative method of computing item-total correlations would probably yield a lower, but likely still impressive, value.

As indicated in Table 2, five projects have evaluated the internal consistency of the AUDIT using Cronbach’s α’s. Values derived suggest high consistency, thus suggesting that the AUDIT is measuring a construct in a reliable fashion. Factor analysis of the AUDIT in a sample of alcoholics and substance abusers in treatment revealed a single principal component with all items loading quite highly on it.¹⁶ Although in developing the AUDIT, the authors selected items to measure the three conceptually distinct dimensions of intake, dependence, and adverse consequences of drinking; these hypothetical dimensions and the individual items that reflect them seem to be highly related.

Because alternative forms of the AUDIT do not exist, parallel form reliability cannot be computed. Neither did we find projects dealing with split-half reliability, a phenomenon perhaps due to the small number of items on the test. Also, as with Cronbach α’s, such reliability studies might be slightly misleading for a test primarily aimed at achieving predictive validity, even at the cost of having somewhat diminished interitem relationships.

A more troubling omission in the published research literature on the AUDIT, as in that for other alcohol screening tests, is determination of test-retest reliability. Considering that the AUDIT generally covers the entire past year, such studies are fully reasonable and might, perhaps, use a test-retest interval of at least a couple of weeks. Most important is determination of test-retest reliability when the AUDIT is dichotomously scored as positive or negative.

### Table 2. Coefficient α’s Reported for the AUDIT

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Subjects</th>
<th>Nature of S’s</th>
<th>Cut-off</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>82 Substance abusers</td>
<td>0.94</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>136 Alcoholics without substance abuse in treatment</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>132 Primary care patients</td>
<td>0.77</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>832 Individuals arrested for driving while intoxicated</td>
<td>0.83</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>204 College students</td>
<td>0.80</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>287 Primary care patients over age 18</td>
<td>0.85</td>
<td></td>
</tr>
</tbody>
</table>

### REMAINING RESEARCH QUESTIONS

Despite the high level of research activity on the AUDIT, several issues are as yet unresolved.

As with many self-report alcohol screening measures, the AUDIT begins with items dealing with quantity, frequency, and intensity of drinking. In a thought-provoking, but unreplicated, study, Steinweg and Worth²⁸ found that directly asking about amount and frequency of drinking before giving the CAGE decreased sensitivity of the scale to 0.32 vs 0.95 when the questions were not preceded by directed queries on drinking. It would be worthwhile to determine possible effects on accuracy of the AUDIT that might occur if the items were reordered.

Although the test developers recommend that the AUDIT could be embedded within a broader general health risk screening instrument, no research project has concerned whether the validities and other psychometric properties of the AUDIT would thereby be changed, as opposed to the AUDIT being used as an independent scale.

The AUDIT is unique among alcohol screening tests in that it is intended to serve as a criterion valid measure, but also to maintain construct validity. On the basis of statistical analyses, rational consideration, and assumed enhanced validity in screening for alcohol problems, items were chosen to reflect the three dimensions of drinking intake, dependence, and adverse consequences. As noted, one study,¹⁶ however, suggests that the AUDIT is saturated by a single factor. The very high coefficient α’s support such a position as well. Further factor analyses using more heterogeneous samples and attempts to extract varying numbers of factors are needed to substantiate the presence of the three intended factors and to determine if the factor structure of the test varies as a function of the nature of the sample along dimensions such as severity of alcohol problems, gender, culture, and social stability. That assessment of the three domains improves predictive validity of the test as opposed to simply selecting items based on their relationship with each other and the criterion remains to be demonstrated.

Future studies on the AUDIT might also consider the significance of scores above the cut-off. Although technically speaking, higher scores simply indicate greater likelihood of the patient actually having the trait of interest, such scores may also reflect greater severity of the trait. It would be of interest to determine if higher AUDIT scores are associated with more symptoms or more severe manifestations of alcohol dependence or harmful drinking, a possibility suggested by Donovan et al.²⁹

Studies of the AUDIT and alternative alcohol screening measures might also relate the tests singly and in combination to diagnosis. Because there seems to be a sizable group of patients who score positively on only one of the self-report screening tests, it may well be that, as in the case of biochemical markers carbohydrate-deficient transferrin and GGT, using the tests in combination would significantly enhance sensitivity without seriously compromising...
specification. Interestingly, the Skipsey et al. project determined that, whereas the areas under the receiver operating curves for a current diagnosis of alcohol abuse/dependence or current hazardous drinking were consistently higher for the AUDIT than for the MAST, at more extreme scores on either test concordance between results as positive or negative was higher.

A major issue in the discussion of a research agenda for screening tests, such as the AUDIT, is development of a comprehensive methodology that permits unequivocal statements to be made about accuracy, feasibility, and utility. Validation procedures for self-report screening tests have varied in sophistication, with some studies providing little more than comparison of one self-report screening procedure, such as the CAGE or MAST, with another self-report instrument, such as the Diagnostic Interview Schedule, which contains many of the same items. In the absence of a commonly accepted “gold standard,” it is recommended that a psychological test validation approach be used, one that examines construct validity, discriminative validity, and predictive validity. One substitute for the gold standard criterion is what has come to be called the LEAD standard. This acronym stands for longitudinal evaluation using all available data, meaning that a complete diagnostic evaluation is conducted on all subjects screened using semistructured diagnostic interviews, laboratory data, interviews with collaterals, and medical records to arrive at the criterion diagnosis against which the screening test is evaluated.

CONCLUSIONS

Beyond its potential for clinical applications, the AUDIT represents a new approach to the development and validation of alcohol screening procedures. This approach attempts to shift the focus from alcoholism as a clinical entity to a public health perspective that emphasizes the early detection of a broader range of alcohol-related problems, only one of which is chronic alcoholism. Screening is used not only for case finding, as the MAST and CAGE have been traditionally used, but also for detection of risk factors, such as hazardous levels of alcohol consumption. In the latter situation, screening would lead to health education rather than referral to formal treatment.

As a screening test for alcohol problems, the AUDIT seems to be quite sensitive and specific. Higher scores on the measure seem associated both with greater likelihood of the condition and probably with greater severity of it. In direct contrasts with other self-report alcohol screening tests, the AUDIT generally exhibits validity at least equal to that of the other measure and in some instances higher. Although we believe that further investigation is needed on the test, similar research recommendations would pertain as well to alternative self-report or biochemical screening tests for alcohol problems. Brevity and ease of administration of the test, absence of copyright fee, and the rich multinational database used in developing the AUDIT argue further for its applicability across a wide range of contexts and populations at risk.

APPENDIX: THE AUDIT QUESTIONNAIRE

1. How often do you have a drink containing alcohol?
   - Never
   - Monthly or less
   - 2–4 times/month
   - 2–3 times/week
   - 4+ times/week

2. How many drinks containing alcohol do you have on a typical day when you are drinking?
   - 1 or 2
   - 3 or 4
   - 5 or 6
   - 7 or 9
   - 10 or more

3. How often do you have six or more drinks on one occasion?
   - Never
   - Less than monthly
   - Monthly
   - Weekly
   - Daily or almost daily

4. How often during the last year have you found that you were not able to stop drinking once you had started?
   - Never
   - Less than monthly
   - Monthly
   - Weekly
   - Daily or almost daily

5. How often during the last year have you failed to do what was normally expected of you because of drinking?
   - Never
   - Less than monthly
   - Monthly
   - Weekly
   - Daily or almost daily

6. How often during the last year have you needed a first drink in the morning to get yourself going after a heavy drinking session?
   - Never
   - Less than monthly
   - Monthly
   - Weekly
   - Daily or almost daily

7. How often during the last year have you had a feeling of guilt or remorse after drinking?
   - Never
   - Less than monthly
   - Monthly
   - Weekly
   - Daily or almost daily

8. How often during the last year have you been unable to remember what happened the night before because you had been drinking?
   - Never
   - Less than monthly
   - Monthly
   - Weekly
   - Daily or almost daily

9. Have you or someone else been injured as a result of your drinking?
   - No
   - Yes, but not in the last year
   - Yes, during the last year

10. Has a relative, friend, doctor, or other health worker been concerned about your drinking or suggested that you should cut down?
    - No
    - Yes, but not in the last year
    - Yes, during the last year

The AUDIT is scored by summing the weights associated with the response selected for each item.
REFERENCES


4. MacAndrew C: The differentiation of male alcoholic outpatient from non-alcoholic psychiatric outpatients by means of the MMPI. Q J Stud Alcohol 26:238–246, 1965


