HHS Public Access
Author manuscript
Assessment. Author manuscript; available in PMC 2020 September 01.

Published in final edited form as:
Assessment. 2020 September ; 27(6): 1089–1099. doi:10.1177/1073191118792091.

Identifying At-Risk College Student Drinkers With the AUDIT-US: A Receiver Operating Characteristic Curve Analysis

Michael B. Madson[1], **Joshua W. Schutts**[2], **Hallie R. Jordan**[1], **Margo C. Villarosa-Hurlocker**[1,3], **Robert B. Whitley**[1], **Richard S. Mohn**[1]

[1]The University of Southern Mississippi, Hattiesburg, MS, USA

[2]University of West Florida, Pensacola, FL, USA

[3]Center on Alcoholism, Substance Abuse, and Addictions, Albuquerque, NM, USA

Abstract

The Alcohol Use Disorders Identification Test (AUDIT) is the gold standard screening measure. Recently, there has been increasing call to update the measure to reflect harmful drinking standards in the United States. The purpose of this study was to use receiver operating characteristic curve analysis to evaluate the AUDIT and the United States version (AUDIT-US). Participants were 382 traditional age (M = 20.2, SD = 1.5) college students (68.7% female, 64.9% White) who had consumed alcohol at least once in the 30 days prior to participating. Although results provide evidence for the AUDIT and the AUDIT-US as valid screening tools, the Consumption subscale of the AUDIT-US performed the best in predicting at-risk college drinkers. The Consumption subscale of the AUDIT-US with a single cutoff score of four appears to be the optimal and most parsimonious method of identifying at-risk college drinkers.

Keywords

at-risk drinking; college students; screening; Alcohol Use Disorders Identification Test

Identifying hazardous drinking remains an important health care issue in the United States mainly to avert the development of alcohol use disorders (AUDs). Successful efforts to implement programs such as the Screening and Brief Intervention and Referral to Treatment in health care settings provide growing evidence for the need to identify and address hazardous drinking (Babor, Del Boca, & Bray, 2017). One population with some of the highest rates of hazardous drinking are college students, who drink even more than their noncollege peers (Johnston, O'Malley, Bachman, Schulenberg, & Miech, 2016). In fact, the majority of college students drink and roughly half engage in hazardous drinking behaviors (Johnston et al., 2016). Hazardous alcohol use has been a target for prevention and intervention on college campuses given the associated academic, personal, physical, and

Corresponding Author: Michael B. Madson, Department of Psychology, The University of Southern Mississippi, 118 College Drive, #5025, Hattiesburg, MS 39406-5025, USA., michael.madson@usm.edu.

Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

social consequences frequently experienced when drinking at these levels (White & Hingson, 2013). Even though students who engage in hazardous drinking experience alcohol-related negative consequences, they rarely seek help for alcohol use problems (~5%; Blanco et al., 2008). Therefore, the identification of students at risk for hazardous alcohol use is important for indicated prevention and intervention programs (Barry, Chaney, Stellefson, & Dodd, 2015) and identifying those at risk for developing an AUD (Hagman, 2016). Screening is one method to identify at-risk alcohol use and relies on validated tools that can effectively identify those who may or may not be at-risk. The Alcohol Use Disorders Identification Test (AUDIT; Saunders, Aasland, Babor, De La Fuente, & Grant, 1993) is the gold standard screening tool.

Originally developed in the health care setting, the 10-item AUDIT was designed to serve as an international/cross-cultural screening tool for hazardous drinking and those at risk for developing an AUD. The scale assesses alcohol consumption (three items; AUDIT–Consumption [AUDIT-C]), alcohol-related problems (two items), and adverse psychological reactions (five items; Saunders et al., 1993). The empirical evidence for the AUDIT is strong across a range of populations and in a variety of settings (Bohn, Babor, & Kranzler, 1995). Among college students, there is sound psychometric evidence for the AUDIT in identifying at-risk drinkers and predicting alcohol use problems (Fleming, Barry, & MacDonald, 1991; Kokotailo et al., 2004). Furthermore, the AUDIT is the most published measure of hazardous drinking and alcohol-related negative consequences in the college student literature (Devos-Comby & Lange, 2008). A key feature of the AUDIT is the use of empirically derived cutoff scores for identifying hazardous drinking risk and potential of developing an AUD.

Research has evolved on which cutff scores are best for identifying college students at risk for hazardous drinking and alcohol use problems. Differing cutoff scores ranging from 6 to 15 out of 40 are based on different sensitivity and specificity scores at that cutoff within varying samples (e.g., Fleming et al., 1991; Hagman, 2016; Kokotailo et al., 2004; Saunders et al., 1993). Additionally, researchers have found differing cutoff scores when using the AUDIT and the AUDIT-C (Barry et al., 2015; DeMartini & Carey, 2012), such that the AUDIT-C cutoff scores tend to outperform the AUDIT total cutoff scores as a screener. Furthermore, there is a call for developing gender-specific cutoff scores as at-risk drinking is defined differently for males and females (DeMartini & Carey, 2012; Olthuis, Zamboanga, Ham, & Van Tyne, 2011). Initial research using receiver operating characteristic (ROC) curve analysis has identified AUDIT-C cutoff scores of 5 for females and 7 for males (DeMartini & Carey, 2012). However, Hagman (2015) found an optimal score of 3 for females and 5 for males. Thus, there remains an inconsistency in determining the optimal cutoff score. From a clinical utility perspective, these inconsistencies can increase confusion and variability in identifying at-risk students. Thus, utilization of this screening tool might be enhanced through identification of single consistent cutoff score that can be used for men and women.

Additionally, there have been many changes to the conceptualization of risky alcohol use since the publication of the AUDIT almost 30 years ago. In fact, Babor, Higgins-Biddle, and Robaina (2016) recently adapted the AUDIT to be more consistent with the U.S. standards for risky drinking. Specifically, the National Institute on Alcohol Abuse and Alcoholism

(NIAAA) has defined daily and weekly consumption limits to identify “at-risk” drinking as well as more specifically defining heavy episodic or “binge drinking” for men and women. For women, these limits include no more than 3 standard drinks (standard drink = 12 fluid oz of regular beer, 8–9 oz of malt liquor, 5 oz of wine, 1.5 oz shot of distilled spirits) on any day or 7 drinks in a week while the limits include no more than 4 drinks on any day or 14 drinks per week for men (NIAAA, 2010). Similarly, these guidelines define heavy episodic drinking as 5 standard drinks (male) 4 drinks (female) within a 2-hour period (NIAAA, 2015). However, the four response options on the first two items of the AUDIT were limited in their specificity to assess the changes reflected in the NIAAA guidelines. Furthermore, the third item—*How often do you have 6 or more drinks on one occasion*—intended to assess heavy episodic drinking exceeds the NIAAA definition for heavy episodic drinking in the United States. For these reasons, the clinical utility of the AUDIT may be enhanced by better reflecting the safe drinking guidelines in the United States and potentially reducing the potential for false negatives.

The brevity of the AUDIT and AUDIT-C combined with its evidence base lends itself well to various screening environments on college campuses such as health clinics, screening days, counseling centers, and alcohol and drug programs. Yet there is a call from the Centers for Disease Control and Prevention (CDC) to use gender-specific recommendations for harmful drinking as outlined by the NIAAA (2015). As seen in Table 1, changes to the AUDIT and the United States version (AUDIT-US) include expanded response options from 5 to 7 for the first three items (Consumption scale), while the first two questions remain the same. The expanded response options allow for more precise responding. More specifically, responses on Item 1 of the AUDIT range from “never” to “4 or more times a week,” whereas on the AUDIT-US the responses range from “never” to “daily.” Item 3 has been revised from “How often do you have 6 or more drinks on one occasion?” to “How often do you have 5 drinks (male), 4 drinks (female), or more on one occasion?” better representing the definition of heavy episodic drinking in the United States.

It is speculated that the changes recommended by the CDC will better match the AUDIT with U.S. safe drinking guidelines that were not captured in the initial AUDIT. However, in searching the literature, we found no studies that have empirically evaluated the AUDIT-US. Therefore, the present study sought to evaluate the utility of the AUDIT-US in identifying at-risk drinkers using ROC curve analysis. Through ROC curve analysis, we can gauge the specificity and sensitivity of the AUDIT/AUDIT-US to evaluate the performance of these measures. By identifying a better performing measure, we can enhance decision making related to alcohol prevention and interventions on college campuses. We sought to replicate the analyses conducted by DeMartini and Carey (2012) in our analyses of the AUDIT and AUDIT-US to determine if there are performance differences in predicting at-risk drinkers as a group and within sex. We predicted that the AUDIT-US, specifically the AUDIT-US-C, would outperform the AUDIT in sensitivity and specificity in identifying at-risk college drinkers. Because the AUDIT-US-C accounts for gender differences, we predicted a single optimal cutoff score would be identified.

Method

Participants and Procedures

Participants were 382 traditional age college student drinkers (*M* = 20.2, *SD* = 1.5) at a university in the Southern region of the United States. Most participants were female (68.7%) and identified as White (64.9%) or African American (28.1%) with the remaining 7% identifying as other. Participants were recruited through the Psychology Department research management system (SONA system), listserve e-mails, and direct e-mails. Those who participated in the psychology classes received credit as partial fulfillment of a research requirement, while those not in psychology classes were entered into a drawing to receive 1 of 10 ($10) university gift cards. Data were screened to ensure no participant took the survey more than once. After completing an institutional review board-approved informed consent form, participants completed study measures that were presented in random order using a secure online survey system (Qualtrics).

Measures

At-Risk Drinkers.—At-risk drinkers were identified using three questions for females and males based on the NIAAA (2010) guidelines for safe drinking. For females, these questions included the following: In the past year, how many times have you had (a) 4 or more drinks in 2 hours, (b) 3 or more drinks in a day, and (c) 7 or more drinks in a week. For males, the questions included the following: In the past year, how many times have you had (a) 5 or more drinks in 2 hours, (b) 4 or more drinks in a day, and (c) 14 or more drinks in a week. Participants were identified as at-risk drinkers if they indicated engaging in any of these drinking behaviors at least once and were coded as 1 “at-risk” or 0 “not at-risk.”

Alcohol Use Disorders Identification Test.—The 10-item AUDIT (Saunders et al., 1993) assesses hazardous drinking patterns (e.g., “How often do you have 6 or more drinks on one occasion?”). The first three items of the AUDIT capture the quantity and frequency of alcohol use (AUDIT-C). The AUDIT is the leading instrument for the detection of early phase risky drinking patterns across different cultures and age groups (e.g., Reinert & Allen, 2007). The internal consistency estimates in this study were acceptable: AUDIT (.80), AUDIT-C (.77)

Alcohol Use Disorders Identification Test–US.—The 10-item AUDIT-US is a modified version of the original AUDIT that reflects the NIAAA (2010) guidelines for harmful to low-risk drinking (CDC, 2014). As seen in Table 1, the fundamental difference between the AUDIT and AUDIT-US scales is that the response options for Items 1 to 3 on the AUDIT-US were modified, while Items 4 to 10 remain the same. The wording for Item 3 better reflects the U.S. definition of heavy episodic drinking for men and women. The internal consistency estimates in this study were also acceptable: AUDIT-US (.79) and AUDIT-US-C (.78).

Data Analytic Approach

A series of *t* tests were used to test for mean differences between groups of participants. The two groups explored were (a) at-risk versus not at-risk based on the NIAAA (2010) criteria and (b) males versus females. The criteria for statistical significance was adjusted to .01 to control for family wise error. Cohen's *d* was also computed as a measure of effect size in these analyses. Next, we used ROC curve analysis to compare the AUDIT and AUDIT-US scales. The NIAAA (2010) reference standard of at-risk drinking was used to classify participants. ROC curves plot sensitivity (true positive rate) versus 1–specificity (false positive rate) for at-risk status. The primary statistic associated with ROC curves is the area under the curve (AUC). The value of the AUC statistic is the probability that a person identified "at-risk" by the NIAAA (2010) criteria has a greater scale score (e.g., AUDIT, AUDIT-US) than a person identified as "not at-risk." Hosmer, Lemeshow, and Sturdivant (2013) noted that values above .90 are "outstanding," values above .80 are "excellent" and values above .70 are "acceptable." AUC statistics were compared using Hanley and McNeil's (1983) *z* test.

To determine the optimal cutoff score for the various scales, we generated tables of classification statistics. These statistics included sensitivity, specificity, positive (PPV) and negative (NPV) predictive values, the Youden *J* index (Youden, 1950), and the overall diagnostic odds ratio (DOR). The values for PPV and NPV, respectively, represent the percent of positive results ("at-risk") and negative results ("not at-risk") that are true positives (PPV) when at a given cutoff score. Higher values of PPV and NPV give greater confidence that "at-risk" and "not at-risk" test results are actually classified as such using the NIAAA criterion. DOR is calculated as the quotient of the positive (LR+) and negative (LR −) likelihood ratios. The interpretation of DOR is the odds for "at-risk" identification using a particular cutoff score among those identified "at-risk" compared with the odds for "at-risk" identification among those identified "not at-risk."

As observed by DeMartini and Carey (2012), the choice of an optimal cutoff score ideally considers prevalence in the sample and the costs of wrong decisions. For this study, we used the maximum *J* statistic as the primary measure to evaluate optimal cutoff scores. The *J* statistic (Youden, 1950) represents the height of the ROC curve above the chance diagonal at a given cutoff score and gives equal weight to both types of wrong decisions: false negatives and false positives. In cases where the *J* statistic might suggest several optimal cutoff scores, we made a conceptual judgement that false negatives (i.e., Type II error—the test incorrectly identifying an individual as "not at-risk") were worse than false positives (i.e., Type I error —the test incorrectly identifying an individual as "at-risk"). More specifically, if the test missed a person who was actually "at-risk" (an act of omission), that was worse than if the test suggested a person was "at-risk" when they actually were not (an act of commission). Providing an unnecessary intervention—although associated with some costs—was viewed as the lesser evil to ignoring an opportunity to intervene when the person could have benefitted from such. In cases where the optimal *J* statistic was inconclusive, we sought a good balance among the PPV, NPV, and false negative rate, 1–sensitivity (Hagman, 2016).

Results

The median score for self-reported participant consumption was 4 days in the past month (interquartile range = 5). The mean scores for the scales/subscales were 6 for AUDIT (*SD* = 5), 8 for AUDIT-US (*SD* = 6), 4 for AUDIT-C (*SD* = 2), and 6 for AUDIT-US-C (*SD* = 3).

In the overall sample, the prevalence of "at-risk" status was 82.4% (n = 290 of 352). Most participants reported exceeding the NIAAA safe drinking guidelines of 4/5 (female/male) or more drinks within 2 hours (84.7%), 3/4 or more drinks in 1 day (85.2%), and 7/14 or more drinks in a week (69.3%). Of the 290 at-risk participants, 92 (31.7%) were male and 198 (68.3%) were female. In the male subsample, the prevalence of risk was 80.7% (n = 92 of 114). In the female subsample, the prevalence of risk was 83.2% (n = 198 of 238).

We conducted a series of t tests to determine if group differences existed based on at-risk versus not at-risk classification and gender. As expected, the mean scores for the "at-risk" group were significantly higher across all versions of the AUDIT when compared with the "not at-risk" group. The largest difference was observed in the AUDIT-US-C, mean difference = 4, $t(126.35) = 12.62$, $p < .001$, $d = 1.54$. In the AUDIT-US, Cohen's d was 1.13. In the AUDIT, Cohen's d was 0.87, and the AUDIT-C was 1.20. This finding suggests a greater relative difference when using the U.S. versions of the AUDIT scale. We also compared the scores on the several AUDIT versions based on gender. When disaggregating the data set by male/female at-risk drinkers, we found that males had significantly higher scores than females. The largest difference was observed in the AUDIT-US scale, mean difference = 4, $t(120.12) = 4.76$, $p < .001$, $d = 0.63$, Results are presented in Table 2.

ROC Curve Analyses of Classification Performance

Complete Sample.—In the complete sample, the AUC value for the AUDIT-US scale was .84 (95% confidence interval [CI: .79, .90]). That value exceeded the results of the AUDIT scale (AUC = .79, 95% CI [.73, .86]). Similarly, the AUC value for the AUDIT-US-C (AUC = .89, 95% CI [.85, .94]) scale was higher than the AUDIT-C scale (AUC = .85, 95% CI [.80, .90]). The complete results for all versions of the AUDIT are presented in Table 3. Figure 1 displays the ROC curves for each version of the AUDIT. Tables 4 and 5 display, respectively, a range of cutoff scores on the AUDIT-US and AUDIT-US-C. The optimal cutoff score is shaded for effect. In the interest of parsimony, we also compared the classification results for the AUDIT-US and the AUDIT-US-C at their respective optimal cutoff scores. The AUDIT-US-C (at cutoff = 4) is preferable over the AUDIT-US (at cutoff = 6) for three reasons: (a) a higher AUC statistic, .89 versus .84; (b) a higher J statistic, .66 versus .56; and (c) a lower false negative rate, .13 versus .28.

Comparisons by Gender.—The last aim of the study was to compare the AUDIT versions within each gender to determine the best identifier of at-risk drinking. For at-risk males, the AUC value for the AUDIT-US scale was .85 (95% CI [.75, .95]), which was higher than the value for the AUDIT scale (AUC = .81, 95% CI [.70, .93]). For at-risk females, higher AUC values were also found in the AUDIT-US scale (AUC = .85, 95% CI [.79, .91]), versus the AUDIT scale (AUC = .79, 95% CI [.71, .87]). Regardless of gender, the "U.S." version of the AUDIT-C was a better classifier than the original version of the AUDIT-C. For males, the results were as follows: AUDIT-US-C, AUC = .89, 95% CI [.80, .99] versus AUDIT-C, AUC = .88, 95% CI [.78, .98]. For females, the results were as follows: AUDIT-US-C, AUC = .90, 95% CI [.85, .94] versus AUDIT-C, AUC = .84, 95% CI [.79, .90]. The complete results for all versions of the AUDIT, disaggregated by gender, are presented in Table 6. The ROC curves for each version of the AUDIT are presented in Figure 2 (males) and Figure 3 (females).

As with the full data set, we compared both forms of the "AUDIT-US" scale based on their respective optimal cutoff scores. Tables 7 and 8 display, respectively, a range of cutoff scores on the AUDIT-US and AUDIT-US-C, disaggregated by gender. As before, the optimal cutoff score is shaded for effect. For males, the AUDIT-US-C (at cutoff = 4) is recommended over the AUDIT-US (at cutoff = 3) for three reasons: (a) a higher AUC statistic, .90 versus .85; (b) a higher *J* statistic, .70 versus .61; and (c) a lower false negative rate, .08 versus .12. For females, the results are less conclusive. A case can be made to recommend the AUDIT-US-C (at cutoff = 4) for two reasons: (a) a higher AUC statistic, .89 versus .85 and (b) a higher *J* statistic, .64 versus .54. An alternative case can also be made to recommend the AUDIT-US (at cutoff = 4) over the AUDIT-US-C based on a lower false negative rate, .11 versus .16.

In all analyses, the "U.S." version of the AUDIT scale outperformed the original version of the AUDIT. Because the AUDIT-US was developed to incorporate the newer NIAAA (2010) recommendations, we contend that it has better generalizability in addition to its improved classification performance. Our findings also indicated that the AUDIT-C of both the AUDIT and AUDIT-US outperformed their respective full scale versions in the complete data set and the disaggregated gender data sets. That observation was consistent with the findings by DeMartini and Carey (2012). In the interest of implementing one diagnostic test, *regardless of gender*, results from the full sample and the gender subsamples support the use of the AUDIT-US-C (at a cutoff score of 4) to classify males and females.

Discussion

The purpose of this study was to evaluate the AUDIT and AUDIT-US and their subscales, AUDIT-C and AUDIT-US-C, in a sample of college student drinkers following the procedures outlined by DeMartini and Carey (2012). To the best of our knowledge, this was the first attempt to evaluate the AUDIT-US in this fashion. As such, our findings support the idea of developing valid screening measures that are consistent with the NIAAA guidelines for safe drinking—an original need outlined by the AUDIT development team (Babor, & Robaina, 2016). A secondary aim of this study was to contribute the psychometric evaluation of the AUDIT (AUDIT-US) as a screening tool for college student drinkers.

Based on our results, it appears that the AUDIT and AUDIT-US are valid screening tools for use in college student populations given no statistical differences on the *z* test of the AUC statistic. However, when choosing between these tools, the AUDIT-US better reflects high-risk drinking guidelines as outlined by the NIAAA and it does a better job of reducing false negatives, which is arguably the costlier error (i.e., incorrectly predicting the individual is not at-risk). Similarly, a cutoff score of 6 on the AUDIT-US appears to be the best cutoff score for predicting at-risk drinkers. However, the optimal cutoff scores by gender was 5 for males and 4 for females.

Similar to DeMartini and Carey (2012), we found evidence that the AUDIT-C better identified at-risk drinkers when compared with the full AUDIT. Specifically, for males, our findings are consistent with DeMartini and Carey (2012). Although the AUDIT-C was a better predictor of at-risk female drinking, the z test between the AUDIT and AUDIT-C was not significant, and the optimal cutoff scores in the present study were also lower (3 vs. 5) than DeMartini and Carey (2012). We did find that the AUDIT-US-C outperformed the AUDIT and AUDIT-C—especially when the goal of screening was to minimize false negatives for females and males. Furthermore, the optimal cutoff score for the AUDIT-US-C was 4 for males and females, likely because the measure accounts for gender differences in the heavy episodic drinking question (Question 3). Thus, if the goal of screening places more value on minimizing false negatives over false positives and on accounting for gender differences in low-risk alcohol use, the AUDIT-US-C is likely the strongest measure. There is increasing evidence to suggest that the AUDIT-C is better at detecting at-risk drinkers than the full measure in a variety of settings (Aalto, Alho, Halme, & Seppä, 2009; Dawson, Smith, Saha, Rubinsky, & Grant, 2012) including universities (Barry et al., 2015; DeMartini & Carey, 2012). However, there has been variability about what the optimal cutoff score is for identifying at-risk college drinkers (Barry et al., 2015; DeMartini & Carey, 2012) and different scores exist for females and males. Our results extend those findings and provide initial evidence for the value of the AUDIT-US-C for use with college students in the United States. We found a single cutoff score of 4 was the best and accounted for gender differences. Thus, an additional benefit of the AUDIT-US-C may be the practical fact that users simply have to remember a single cutoff score versus multiple scores when screening for risk.

Our findings, while preliminary, have implications for prevention and intervention initiatives. Most notable is that the AUDIT-US (specifically the AUDIT-C) may be more valuable in identifying at-risk college drinkers. This feature of the AUDIT-US is especially important in cases where the costs of misidentifying an at-risk drinker as “not at-risk” are less than misidentifying a non-at-risk drinker as “at-risk.” More specifically, if a student drinker was not at risk but received an intervention, such as a brief motivational intervention, there are still potential benefits as opposed to a student who is an at-risk drinker not receiving the brief motivational intervention. If the at-risk student was not identified, there was missed opportunity at preventing the development of a more severe problem. Thus, there is value for erring on the side of caution when screening for at-risk drinkers on college campuses because the cost is low while still having benefits. Ultimately, the AUDIT-US shows better practical value for that purpose than the AUDIT. An additional clinical implication of our findings is the support for the AUDIT-US-C and the single cutoff score regardless of gender. Having to only use 3 versus 10 items to identify at-risk drinkers allows for parsimony and better clinical utility. Thus, the AUDIT-US-C could be administered in various places on campus such as counseling and health centers or during prevention initiatives such as National College Alcohol Awareness week as a quick but informative screen for at-risk drinking.

Although these results support the use of the AUDIT-US, specifically the AUDIT-C as a screening measure for at-risk college students, study limitations call for caution. Data were collected in one sample at a single university. Replication via a multisite study across the United States would be very valuable. Similarly, these data were collected with a convenience, nontreatment seeking sample and it would be valuable to collect data from individuals seeking health and counseling services, or with individuals participating in alcohol prevention/intervention programs. Additionally, although our sample reflects the university population, the sample is predominantly White non-Hispanic and female. Because differences exist between races in college student drinking it would be important to study the AUDIT-US with more diverse samples. This study only looked at students in college. To advance research on the AUDIT-US, it will be important to study it with noncollege students, adolescents, and adults as well as in diverse settings such as medical facilities, hospitals, and alcohol and drug treatment facilities. Finally, we used the NIAAA guidelines for safe drinking to identify at-risk drinkers given the higher prevalence of hazardous drinking than AUD among college students. While a common method for identifying at-risk drinkers, these guidelines do not represent diagnostic criteria limiting the findings. Thus, it will be important to replicate this study using *Diagnostic and Statistical Manual of Mental Disorders–Fifth edition* (American Psychiatric Association, 2013) diagnostic criteria for AUD.

This study, the first to our knowledge, provided preliminary support for the AUDIT-US as a valid and practical screening tool in identifying at-risk college student drinkers. The AUDIT and AUDIT-US performed well in identifying at-risk drinkers, but to better reflect safe drinking guidelines, reduce false negatives, and to enhance clinical utility, the AUDIT-US, specifically the AUDIT-US-C appears to be a preferable screening tool.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Margo C. Villarosa-Hurlocker's work on this project was supported in part by a training grant (T32-AA018108) from the National Institute on Alcohol Abuse and Alcoholism (NIAAA) in the United States.

References

Aalto M, Alho H, Halme JT, & Seppä K (2009). AUDIT and its abbreviated versions in detecting heavy and binge drinking in a general population survey. Drug and Alcohol Dependence, 103, 25–29. doi:10.1016/j.drugalcdep.2009.02.013 [PubMed: 19395203]

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed). Washington, DC Author.

Babor TF, Del Boca F, & Bray JW (2017). Screening, Brief Intervention and Referral to Treatment: Implications of SAMHSA's SBIRT initiative for substance abuse policy and practice. Addiction, 112, 110–117. doi:10.1111/add.13675 [PubMed: 28074569]

Babor TF, Higgins-Biddle JC, & Robaina K (2016). The Alcohol Use Disorders Identification Test, adapted for use in the United States: A guide for primary care practitioners. Retrieved from http://www.ct.gov/dmhas/lib/dmhas/publications/USAUDIT-2017.pdf

Babor TF, & Robaina K (2016). The Alcohol Use Disorders Identification Test (AUDIT): A review of graded severity algorithms and national adaptations. International Journal of Alcohol and Drug Research, 5, 17–24.

Barry AE, Chaney BH, Stellefson ML, & Dodd V (2015). Evaluating the psychometric properties of the AUDIT-C among college students. Journal of Substance Use, 20, 1–5.

Blanco C, Okuda M, Wright C, Hasin DS, Grant BF, Liu S, & Olfson M (2008). Mental health of college students and their non-college-attending peers: Results from the National Epidemiologic Study on Alcohol and Related Conditions. Archives of General Psychiatry, 65, 1429–1437. doi:10.1001/archpsyc.65.12.1429 [PubMed: 19047530]

Bohn MJ, Babor TF, & Kranzler HR (1995). The Alcohol Use Disorders Identification Test (AUDIT): Validation of a screening instrument for use in medical settings. Journal of Studies on Alcohol, 56, 423–432. doi:10.15288/jsa.1995.56.423 [PubMed: 7674678]

Centers for Disease Control and Prevention. (2014). Planning and implementing screening and brief intervention for risky alcohol use: A step-by-step guide for primary care practices. Atlanta, GA: Centers for Disease Control and Prevention, National Center on Birth Defects and Developmental Disabilities.

Dawson DA, Smith SM, Saha TD, Rubinsky AD, & Grant BF (2012). Comparative performance of the AUDIT-C in screening for *DSM-IV* and *DSM-5* alcohol use disorders. Drug and Alcohol Dependence, 126, 384–388. doi:10.1016/j.drugalcdep.2012.05.029 [PubMed: 22728044]

DeMartini KS, & Carey KB (2012). Optimizing the use of the AUDIT for alcohol screening in college students. Psychological Assessment, 24, 954–963. doi:10.1037/a0028519 [PubMed: 22612646]

Devos-Comby L, & Lange JE (2008). Standardized measures of alcohol-related problems: A review of their use among college students. Psychology of Addictive Behaviors, 22, 349–361. [PubMed: 18778128]

Fleming MF, Barry KL, & MacDonald R (1991). The Alcohol Use Disorders Identification Test (AUDIT) in a college sample. International Journal of the Addictions, 26, 1173–1185. doi:10.3109/10826089109062153 [PubMed: 1743817]

Hagman BT (2015). Toward efficient screening for *DSM-5* alcohol use disorders in college students: Performance of the Audit-C. Addictive Disorders & Their Treatment, 14, 1–15.

Hagman BT (2016). Performance of the AUDIT in detecting *DSM-5* alcohol use disorders in college students. Substance Use & Misuse, 51, 1521–1528. [PubMed: 27438676]

Hanley JA, & McNeil BJ (1983). A method of comparing area under receiver operating characteristic curves derived from the same cases. Radiology, 148, 839–843. [PubMed: 6878708]

Hosmer DW, Lemeshow S, & Sturdivant RX (2013). Applied logistic regression (3rd ed). New York, NY: Wiley.

Johnston LD, O'Malley PM, Bachman JG, Schulenberg JE, & Miech RA (2016). Monitoring the future: National Survey Results on Drug Use, 1975–2015. Retrieved from http://monitoringthefuture.org/pubs/monographs/mtf-overview2015.pdf

Kokotailo PK, Egan J, Gangnon R, Brown D, Mundt M, & Fleming M (2004). Validity of the Alcohol Use Disorders Identification Test in college students. Alcoholism: Clinical & Experimental Research, 28, 914–920. doi:10.1097/01.ALC.0000128239.87611.F5

National Institute on Alcohol Abuse and Alcoholism. (2010). Rethinking drinking (NIH Publication #09–3770). Retrieved from https://pubs.niaaa.nih.gov/publications/RethinkingDrinking/Rethinking_Drinking.pdf

National Institute on Alcohol Abuse and Alcoholism. (2015). Drinking levels defined. Retrieved from https://www.niaaa.nih.gov/alcohol-health/overview-alcohol-consumption/moderate-binge-drinking

Olthuis J, Zamboanga B, Ham L, & Van Tyne K (2011). The utility of a gender-specific definition of binge drinking on the AUDIT. Journal of American College Health, 59, 239–245. doi:10.1080/07448481.2010.497523 [PubMed: 21308582]

Reinert DF, & Allen JP (2007). The Alcohol Use Disorders Identification Test: An update of research findings. Alcoholism: Clinical & Experimental Research, 31, 185–199.

Saunders JB, Aasland GO, Babor TF, De La Fuente JR, & Grant M (1993). Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption-II. Addiction, 88, 791–804. doi:10.1111/j.1360-0443.1993.tb02093.x. [PubMed: 8329970]

White A, & Hingson R (2013). The burden of alcohol use: Excessive alcohol consumption and related consequences among college students. Alcohol Research: Current Reviews, 35, 201–218. [PubMed: 24881329]

Youden WJ (1950). Index for rating diagnostic tests. Cancer, 3(1), 32–35. [PubMed: 15405679]

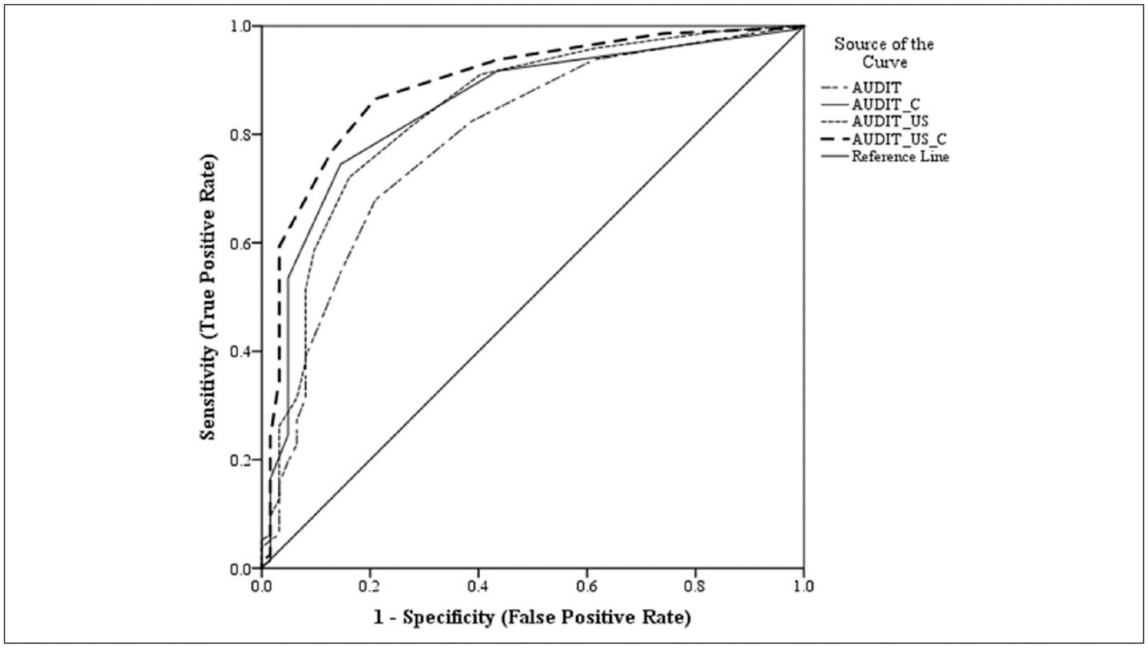

Figure 1.
Complete sample ($n = 290$) ROC curves for all versions of the AUDIT.
Note. ROC = receiver operating characteristic; AUDIT = Alcohol Use Disorders Identification Test.

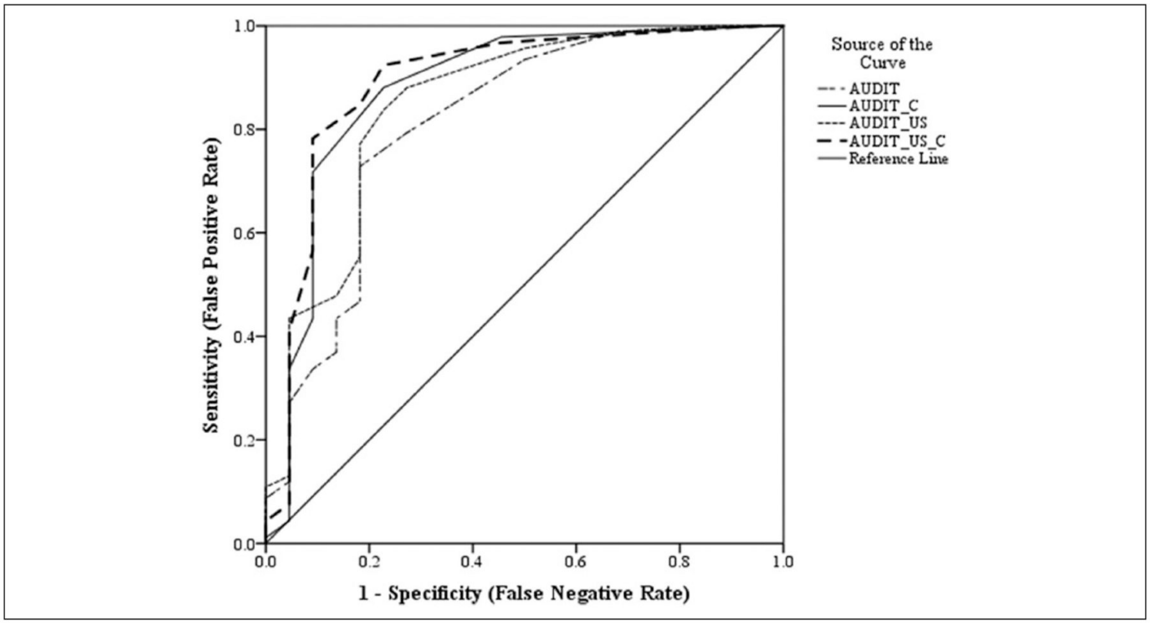

Figure 2.

Male subsample (n = 92) ROC curves for all versions of the AUDIT.

Note. ROC = receiver operating characteristic; AUDIT = Alcohol Use Disorders Identification Test.

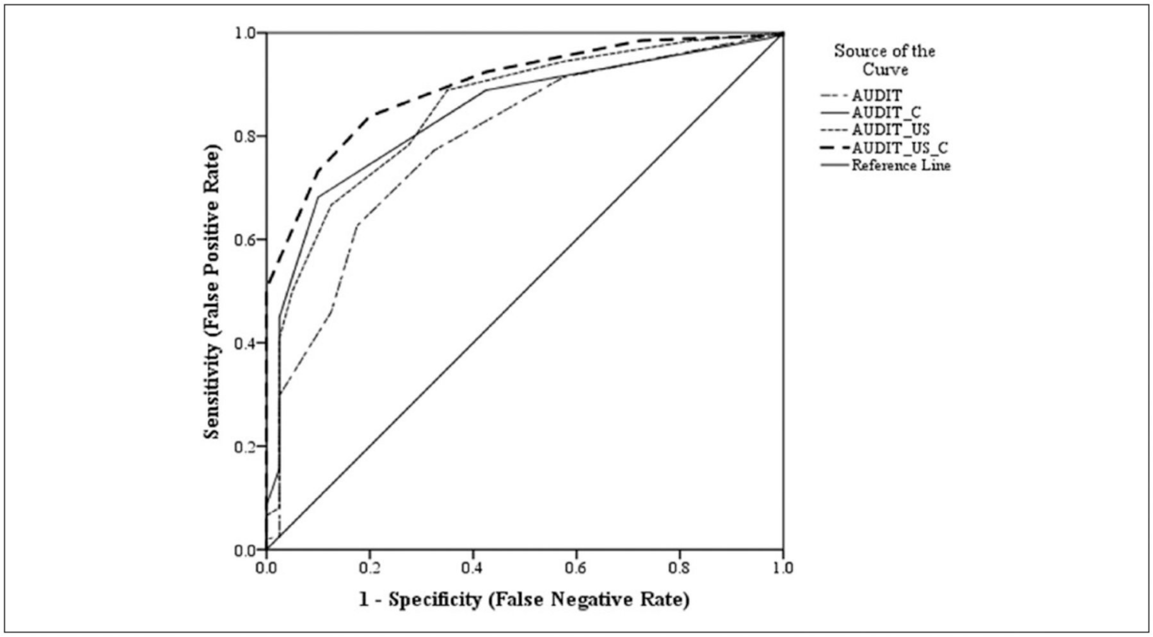

Figure 3.
Female subsample ($n = 198$) ROC curve for all versions of the AUDIT.
Note. ROC = receiver operating characteristic; AUDIT = Alcohol Use Disorders Identification Test.

Table 1.

Changes Between AUDIT and AUDIT-US.

		0	1	2	3	4	5	6
1.	How often do you have a drink containing alcohol?							
	AUDIT	Never	Monthly or less	2–4 Times per month	2–3 Times a week	4 or More times a week		
	AUDIT-US	Never	Less than monthly	Monthly	Weekly	2–3 Times a week	4–6 Times a week	Daily
2.	How many drinks containing alcohol do you have on a typical day when you are drinking?							
	AUDIT	1 or 2	3 or 4	5 or 6	7 to 9	10 or more		
	AUDIT-US	1	2	3	4	5–6	7–9	10 or more
3.	How often do you have 6 or more drinks on one occasion? (AUDIT)							
	How often do you have 5 drinks (male) 4 drinks (female) or more on one occasion? (AUDIT-US)							
	AUDIT	Never	Less than monthly	Monthly	Weekly	Daily or almost daily		
	AUDIT-US	Never	Less than monthly	Monthly	Weekly	2–3 Times a week	4–6 Times a week	Daily

Note. AUDIT = Alcohol Use Disorders Identification Test; AUDIT-US = AUDIT–United States version.

Table 2.

Average Scores for At-Risk Drinker Categorization and Gender.

	Full scales		Consumption subscales	
	AUDIT	**AUDIT-US**	**AUDIT**	**AUDIT-US**
	M (SD)	***M (SD)***	***M (SD)***	***M (SD)***
Full sample				
NIAAA criterion				
No (n = 62)	3 (3)	4 (3)	2 (2)	3 (2)
Yes (n = 290)	7 (5)	9 (6)	4 (2)	7 (3)
Welch's t (d)	7.17 (0.87)	9.54 (1.13)	9.75 (1.20)	12.62 (1.54)
p	<.0001	<.0001	<.0001	<.0001
At-risk, by gender				
NIAAA criterion				
Males (n = 92)	9 (6)	12 (6)	5 (3)	8 (3)
Females (n = 198)	6 (4)	8 (5)	4 (2)	6 (3)
Welch's t (d)	4.77 (0.63)	4.76 (0.63)	6.17 (0.83)	5.57 (0.74)
p	<.0001	<.0001	<.0001	<.0001

Note. AUDIT = Alcohol Use Disorders Identification Test; AUDIT-US = AUDIT–United States version; NIAAA = National Institute on Alcohol Abuse and Alcoholism.

Table 3.

Overall Evaluation of Classification Performance for the Complete Sample.

Scale	AUC statistic	Cutoff score	*J*	Sensitivity	Specificity	FPR	FNR	PPV	NPV	DOR
AUDIT	.79	4	.47	.68	.79	.21	.32	.94	.34	7.96
AUDIT-US	.84	6	.56	.72	.84	.16	.28	.95	.39	13.47
AUDIT-C	.85	3	.60	.75	.86	.14	.26	.96	.42	17.23
AUDIT-US-C	.89	4	.66	.87	.79	.21	.13	.95	.54	24.31

Note. AUC = area under the curve; *J* = Youden's statistic; FPR = false positive rate; FNR = false negative rate; PPV = positive predictive value; NPV = negative predictive value; DOR = diagnostic odds ratio; AUDIT = Alcohol Use Disorders Identification Test; AUDIT-US = AUDIT–United States version; AUDIT-C = AUDIT–Consumption; AUDIT-US-C = AUDIT–United States version–Consumption. Optimal cutoff scores are interpreted as "greater than or equal to" a given value.

Table 4.

AUDIT-US Cutoff Analysis in the Complete Sample.

Score	Sensitivity	Specificity	FPR	FNR	*J*	PPV	NPV	DOR
2	.99	.16	.84	.01	.15	.85	.77	19.00
3	.96	.39	.61	.04	.35	.88	.67	14.77
4	.91	.60	.40	.09	.51	.91	.59	14.98
5	.81	.73	.27	.19	.54	.93	.45	11.60
6	.72	.84	.16	.28	.56	.95	.39	13.47
7	.59	.90	.10	.41	.49	.97	.32	13.18
8	.51	.92	.08	.49	.43	.97	.29	12.00
9	.45	.92	.08	.56	.36	.96	.26	9.10
10	.38	.92	.08	.62	.30	.96	.24	6.92

Note. AUDIT-US = AUDIT–United States version; FPR = false positive rate (1–specificity); FNR = false negative rate (1–sensitivity); *J* = Youden's statistic; PPV = positive predictive value; NPV = negative predictive value; DOR = diagnostic odds ratio. Optimal cutoff scores are shaded in gray and interpreted as "greater than or equal to" a given value.

Table 5.

AUDIT-US-C Cutoff Analysis in the Complete Sample.

Score	Sensitivity	Specificity	FPR	FNR	*J*	PPV	NPV	DOR
1	1.00	.00	.98	.00	.01	.84	.51	5.40
2	.99	.02	.74	.01	.24	.87	.78	24.49
3	.94	.26	.44	.06	.50	.92	.64	19.65
4	.87	.57	.21	.13	.66	.95	.54	24.31
5	.77	.79	.13	.23	.64	.97	.43	22.48
6	.59	.87	.03	.41	.56	.99	.32	44.07
7	.47	.97	.03	.53	.44	.99	.26	27.04
8	.35	.97	.03	.66	.31	.98	.22	15.93
9	.25	.97	.02	.76	.23	.99	.20	19.96

Note. AUDIT-US-C = AUDIT–United States version–Consumption; FPR = false positive rate (1–specificity); FNR = false negative rate (1–sensitivity); *J* = Youden's statistic; PPV = positive predictive value; NPV = negative predictive value; DOR = diagnostic odds ratio. Optimal cutoff scores are shaded in gray and interpreted as "greater than or equal to" a given value.

Table 6.

Overall Evaluation of Classification Performance for the At-Risk Samples of Males and Females.

Scale	AUC statistic	Cutoff score	*J*	Sensitivity	Specificity	FPR	FNR	PPV	NPV	DOR
Males (n = 92)										
AUDIT	.81	5	.55	.73	.82	.18	.27	.94	.42	12.03
AUDIT-US	.85	6	.61	.84	.77	.23	.16	.94	.53	17.49
AUDIT-C	.88	3	.65	.88	.77	.23	.12	.94	.61	24.97
AUDIT-US-C	.89	4	.70	.92	.77	.23	.08	.94	.71	41.40
Females (n = 198)										
AUDIT	.79	4	.45	.63	.83	.18	.37	.95	.31	7.89
AUDIT-US	.85	4	.54	.89	.65	.35	.11	.93	.54	14.87
AUDIT-C	.84	3	.58	.68	.90	.10	.32	.97	.36	19.30
AUDIT-US-C	.90	4	.64	.84	.80	.20	.16	.95	.50	20.69

Note. AUC = area under the curve; AUDIT = Alcohol Use Disorders Identification Test; AUDIT-US = AUDIT–United States version; AUDIT-C = AUDIT–Consumption; AUDIT-US-C = AUDIT–United States version–Consumption; *J* = Youden's statistic; FPR = false positive rate; FNR = false negative rate; PPV = positive predictive value; NPV = negative predictive value; DOR = diagnostic odds ratio. Optimal cutoff scores are interpreted as "greater than or equal to" a given value.

Table 7.

AUDIT-US Cutoff Analysis in the At-Risk Male and Female Samples.

Score	Sensitivity	Specificity	FPR	FNR	*J*	PPV	NPV	DOR
Males (*n* = 92)								
2	1.00	.14	.86	.00	.14	.83	1.00	
3	.99	.32	.68	.01	.31	.86	.87	41.92
4	.96	.50	.50	.04	.46	.89	.74	22.26
5	.88	.73	.27	.12	.61	.93	.59	19.53
6	.84	.77	.23	.16	.61	.94	.53	17.49
7	.77	.82	.18	.23	.59	.95	.46	15.22
8	.74	.82	.18	.26	.56	.94	.43	12.73
9	.62	.82	.18	.38	.44	.93	.34	7.33
10	.55	.82	.18	.45	.37	.93	.30	5.58
Females (*n* = 198)								
2	.99	.18	.83	.02	.16	.86	.70	13.93
3	.94	.43	.58	.06	.37	.89	.61	12.46
4	.89	.65	.35	.11	.54	.93	.54	14.87
5	.78	.73	.28	.22	.51	.93	.40	9.51
6	.67	.88	.13	.33	.54	.96	.35	14.02
7	.50	.95	.05	.50	.45	.98	.28	19.00
8	.41	.98	.03	.59	.38	.99	.25	26.99
9	.36	.98	.03	.64	.34	.99	.24	22.32
10	.30	.98	.03	.70	.27	.98	.22	16.56

Note. AUDIT-US = AUDIT–United States version; FPR = False positive rate (1–specificity); FNR = false negative rate (1–sensitivity); *J* = Youden's statistic; PPV = positive predictive value; NPV = negative predictive value; DOR = diagnostic odds ratio. Optimal cutoff scores are shaded in gray and interpreted as "greater than or equal to" a given value.

Table 8.

AUDIT-US-C Cutoff Analysis in the At-Risk Male and Female Samples.

Score	Sensitivity	Specificity	FPR	FNR	*J*	LR+	LR−	PPV	NPV	DOR
Males ($n = 92$)										
1	1.00	.05	.96	.00	.04	1.05	.00	.81	1.00	
2	.99	.23	.77	.01	.22	1.28	.05	.84	.83	26.40
3	.97	.55	.46	.03	.51	2.13	.06	.90	.80	35.10
4	.92	.77	.23	.08	.70	4.07	.10	.94	.71	41.40
5	.85	.82	.18	.15	.67	4.66	.19	.95	.56	25.07
6	.78	.91	.09	.22	.69	8.6	.24	.97	.50	36.04
7	.69	.91	.09	.32	.59	7.53	.35	.97	.41	21.72
8	.57	.91	.09	.44	.47	6.21	.48	.96	.33	12.97
9	.41	.96	.05	.59	.37	9.18	.61	.97	.28	14.93
Females ($n = 198$)										
1	1.00	.00	1.00	.01	.00	1.00		.83	.00	
2	.99	.28	.73	.02	.26	1.36	.05	.87	.79	24.91
3	.92	.58	.43	.08	.50	2.17	.13	.92	.60	16.45
4	.84	.80	.20	.16	.64	4.19	.20	.95	.50	20.69
5	.73	.90	.10	.27	.63	7.32	.30	.97	.40	24.58
6	.51	1.00	.00	.50	.51		.50	1.00	.29	
7	.37	1.00	.00	.63	.37		.63	1.00	.24	
8	.24	1.00	.00	.76	.24		.76	1.00	.21	
9	.17	1.00	.00	.83	.17		.83	1.00	.20	

Note. AUDIT-US-C = AUDIT–United States version–Consumption; FPR = false positive rate (1–specificity); FNR = false negative rate (1–sensitivity); *J* = Youden's statistic; LR+ = positive likelihood ratio; LR− = negative likelihood ratio; PPV = positive predictive value; NPV = negative predictive value; DOR = diagnostic odds ratio. Optimal cutoff scores are shaded in gray.