

Critical Review

Facilitating Screening and Brief Interventions in Primary Care: A Systematic Review and Meta-Analysis of the AUDIT as an Indicator of Alcohol Use Disorders

Shannon Lange , Kevin Shield, Maristela Monteiro, and Jürgen Rehm 

Background: The Alcohol Use Disorders Identification Test (AUDIT) was developed for use in primary health care settings to identify hazardous and harmful patterns of alcohol consumption, and is often used to screen for alcohol use disorders (AUDs). This study examined the AUDIT as a screening tool for AUDs.

Methods: A systematic literature search was performed of electronic bibliographic databases (CINAHL, Embase, ERIC, MEDLINE, PsycINFO, Scopus, and Web of Science) without language or geographic restrictions for original quantitative studies published before September 1, 2018, that assess the AUDIT's ability to screen for AUDs. Random-effects meta-regression models were constructed by sex to assess the potential determinants of the AUDIT's specificity and sensitivity. From these models and ecological data from the Global Information System on Alcohol and Health, the true- and false-positive and true- and false-negative proportions were determined. The number of people needed to be screened to treat 1 individual with an AUD was estimated for all countries globally where AUD data exist, using a specificity of 0.95.

Results: A total of 36 studies met inclusion criteria for the meta-regression. The AUDIT score cut-point was significantly associated with sensitivity and specificity. Standard drink size was found to affect the sensitivity and specificity of the AUDIT for men, but not among women. The AUDIT performs less well in identifying women compared to men, and countries with a low prevalence of AUDs have higher false-positive rates compared to countries with a higher AUD prevalence.

Conclusions: The AUDIT does not perform well as a screening tool for identifying individuals with an AUD, especially in countries and among populations with a low AUD prevalence (e.g., among women), and thus should not be used for this purpose.

Key Words: Alcohol Use Disorders, Alcohol Use Disorders Identification Test, Classification Accuracy, Screening.

ALCOHOL USE DISORDERS (AUDs) contribute markedly to the burden of disease (Rehm et al., 2013). Specifically, AUDs impact disability, functioning, and mortality risk due to people with an AUD consuming high volumes of alcohol and engaging in detrimental patterns of drinking (Rehm et al., 2003). As a result, on numerous

occasions the World Health Organization (WHO) and other United Nations organizations have called for a reduction in the harmful use of alcohol (World Health Organization, 2018c); however, to date, there has been limited success in achieving such reductions, and rather than reaching reduction targets, the harmful use of alcohol seems to be

Institute for Mental Health Policy Research (SL, KS, JR), Centre for Addiction and Mental Health, Toronto, ON, Canada; Dalla Lana School of Public Health (KS, JR), University of Toronto, Toronto, ON, Canada; Noncommunicable Diseases and Mental Health Department (MM), Pan American Health Organization (PAHO), Washington, District of Columbia; Institute of Medical Science (JR), Faculty of Medicine, University of Toronto, Toronto, ON, Canada; Campbell Family Mental Health Research Institute (JR), Centre for Addiction and Mental Health, Toronto, ON, Canada; Department of Psychiatry (JR), University of Toronto, Toronto, ON, Canada; Institute of Clinical Psychology and Psychotherapy & Center of Clinical Epidemiology and Longitudinal Studies (CELOS) (JR), Technische Universität Dresden, Dresden, Germany; and Department of International Health Projects (JR), Institute for Leadership and Health Management, I.M. Sechenov First Moscow State Medical University, Moscow, Russia.

Received for publication May 1, 2019; accepted July 29, 2019.

Reprint requests: Shannon Lange, PhD, Institute for Mental Health Policy Research, Centre for Addiction and Mental Health, 33 Russell St, Room T521, Toronto, ON M5S 2S1, Canada; Tel.: +1-416-535-8501 ext. 34512; Fax: +1-416-595-6899; E-mails: shannon.lange@camh.ca, shannonlange@hotmail.com

© 2019 The Authors. *Alcoholism: Clinical & Experimental Research* published by Wiley Periodicals, Inc. on behalf of Research Society on Alcoholism.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

DOI: 10.1111/acer.14171

increasing (Manthey et al., 2019). Consequently, the WHO launched the SAFER alcohol control initiative, aimed at preventing and reducing alcohol-related deaths and disability (World Health Organization, 2018b). One of the 5 key alcohol control measures of SAFER is facilitating access to screening, brief interventions, and treatment.

In order to enable brief interventions for hazardous and harmful patterns of alcohol consumption and treatment of AUDs, reliable and valid tools to identify populations in need have to be identified. Although the WHO Alcohol Use Disorders Identification Test (AUDIT; Babor et al., 2001) was primarily developed for the former purpose (Saunders et al., 1993), and is currently the most widely used tool to identify hazardous and harmful drinking (Babor and Robaina, 2016), it is often used to screen for AUDs, and thus has the potential to facilitate the latter purpose.

AUDs can be treated, largely within primary health care settings (Carvalho et al., 2019; Rehm et al., 2016), and thus, the AUDIT may serve as a valuable tool to initiate such interventions, if the AUDIT can accurately identify AUDs. Accordingly, the current study had the following aims:

1. to examine the AUDIT as a screening tool for AUDs, inclusively, and alcohol dependence, independent of other AUDs;
2. to determine the optimal cut-points (the minimum score obtained on the AUDIT that is indicative of an AUD) and determinants (factors that decisively affect the ability of the AUDIT to identify AUDs); and
3. to apply these cut-points to estimate, by sex and country, the number of people needed to be screened to treat 1 individual with an AUD, in order to establish the full clinical utility of the AUDIT.

Given that a diagnosis of an AUD is associated with a high level of stigmatization, over and above the stigma associated with other mental disorders (Schomerus et al., 2011), the specificity of a screening instrument for AUDs should be maximized in order to increase the likelihood that an individual being treated, whether within a primary health care setting or within a specialized health care setting, does, in fact, require such treatment. Thus, the AUDIT cut-points that provide a specificity of 0.95 were considered optimal. A secondary analysis was performed using a specificity of 0.90 and 0.98.

MATERIALS AND METHODS

The systematic literature review and meta-analysis was conducted and reported according to the standards set out in the didactic guidelines developed by Devillé and colleagues (2002), the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (Liberati et al., 2009), and the Meta-analysis of Observational Studies in Epidemiology (Stroup et al., 2000).

Definitions

AUD was defined as: (i) an AUD as per the Diagnostic and Statistical Manual of Mental Disorders (DSM)-5; (ii) alcohol abuse

and alcohol dependence as per the DSM-III or DSM-IV; or (iii) the harmful use of alcohol and alcohol dependence syndrome as per the International Statistical Classification of Diseases and Related Health Problems (ICD)-9 or ICD-10.

Comprehensive Systematic Literature Search

A comprehensive systematic literature search was performed to identify all studies that have validated the AUDIT using the diagnosis of an AUD as the criterion standard. The search was conducted in multiple electronic bibliographic databases, including CINAHL, Embase, ERIC, MEDLINE, PsycINFO, Scopus, and Web of Science. The search was conducted using the following keywords: (i) (“alcohol use disorder* identification test” OR “AUDIT”) OR (“alcohol use disorder*” OR “alcoholism” OR “alcohol abuse” OR “alcohol dependence”) AND “screen*”) AND (ii) “sensitivity and specificity” OR “predictive value*” OR “classification accuracy” OR “receiver operator characteristic analys*” OR “roc analys*” OR “roc curve*.” The search was performed to identify peer-reviewed studies published before September 1, 2018, without language or geographic restrictions. Further, the content pages of the major medical and epidemiological journals, as well as citations in the relevant articles, were manually screened. The full review protocol is available in PROSPERO (<http://www.crd.york.ac.uk/PROSPERO/>), registration number CRD42018110117.

Study Selection and Data Extraction. Study selection began by screening titles and abstracts for inclusion. Then, full-text articles of all studies screened as potentially relevant were considered. Two investigators conducted each study selection step independently; any disagreements were reconciled by team discussion. All data were extracted by 1 investigator and then independently cross-checked by a second investigator for accuracy against the original studies. All discrepancies were reconciled by team discussion. Non-English-language studies deemed to be potentially relevant were translated either by colleagues who were native speakers or fluent in the respective language or by using Google Translate (and subsequently cross-checked by a native or fluent speaker). The following data were extracted from each study deemed relevant: reference, year of study, country, population, setting, sample size, number of cases, sex, AUDIT cut-point, tool(s) used to ascertain the existence of an AUD, the AUD diagnostic criteria used, sensitivity, and specificity.

Eligibility Criteria. Articles were retained if they: (i) consisted of original, quantitative research published in a peer-reviewed journal; (ii) administered the AUDIT; (iii) identified the prevalence of current AUDs (defined as the presence of an AUD, inclusively, in the past 12 months) and/or alcohol dependence, independent of other AUDs, in their study sample: (i) using a valid measure of AUD (National Institute on Alcohol Abuse and Alcoholism, 2003), (ii) as per the diagnosis of a trained psychiatrist using either the DSM or ICD criteria, or (iii) established via the occurrence of alcohol withdrawal syndrome, as an indication of alcohol dependence, by sex; (iv) reported the sensitivity and specificity for the AUDIT with respect to AUDs, inclusively, and alcohol dependence, independent of other AUDs, by sex; and (v) specified the cut-point(s) used to obtain the respective sensitivity and specificity estimates. Articles were excluded if they: (i) used a modified scoring approach (i.e., they did not use the total cumulative scoring approach, as described in Babor and colleagues 2001); (ii) reported sensitivity and specificity for cut-point ranges or categories (e.g., 1 to 3, 4 to 7, 8 to 11, and ≥ 12); or (iii) reported sensitivity and specificity for both sexes combined.

If a study reported the measures of classification accuracy both for the overall sample including abstainers and for current drinkers only, the measures for current drinkers only were retained. In cases where multiple studies used the same dataset or cohort, the study

with the larger sample size was included—that is, studies that reported on the full sample were given priority over studies that reported on subsamples (e.g., specific ethnicities). If a study reported sensitivity and specificity for an AUD by using both: (i) the DSM-IV and DSM-5, the estimates for the DSM-5 were retained as it is the most recent version; or (ii) the DSM-IV and ICD-10, the estimates for the DSM-IV were retained as it has been shown to have better between measure agreement than the ICD-10 when compared with valid measures of substance use disorders, including AUDs (Hasin et al., 2006).

Risk of Bias Assessment. Risk of bias assessment was completed by 2 of the investigators, working independently, and using the revised Quality Assessment of Diagnostic Accuracy Studies tool (QUADAS-2; Whiting et al., 2011).

Statistical Analysis

Cut-point-specific sensitivity and specificity data for the AUDIT, which were used for the screening of AUDs and alcohol dependence, were aggregated using random-effects meta-analyses (DerSimonian and Laird, 1986). These data were transformed using the Freeman–Tukey double arcsine transformation and back-transformed using the harmonic mean (Freeman and Tukey, 1950). Random-effects meta-regression models were then conducted to investigate the role of indicator variables, including country-specific standard drink sizes and cut-points. The standard drink size referenced when administering the AUDIT will impact: (i) the number of drinks consumed by the respondent on a typical drinking day (question 2), and (ii) how often the respondent drank 6 or more drinks on 1 occasion (question 3). As per the AUDIT manual, the respective questions assume a standard drink size of 10 g of ethanol (EtOH), and adjustment should be made to the response categories in order to account for the standard drink size in the country it is being administered in (Babor et al., 2001). Given that the responses to these 2 questions will influence the score obtained on the AUDIT and that average volume of alcohol consumed is a risk factor for an AUD diagnosis (Rehm et al., 2003), standard drink size is likely to affect the cut-point indicative of an AUD. Country-specific drink sizes (measured in grams of EtOH) were obtained from 2 published systematic reviews of the European (Mongan and Long, 2015) and global (Kalinowski and Humphreys, 2016) standard drink sizes used in low-risk drinking guidelines and surveys. The best-fitting model was based on the Akaike information criterion and the Bayesian information criterion, as well as on R-squared statistics. Heterogeneity between double-arcsine-transformed estimates of the sensitivity and specificity of the AUDIT was assessed using the τ^2 statistic (Deeks et al., 2008; Higgins and Thompson, 2002). Publication bias was examined by visually inspecting the funnel plot (standard error plotted against the point estimate) for a skewed distribution, and by employing Egger's weighted regression test for small-study effects (Egger et al., 1997).

For 190 countries where data on the prevalence of AUDs and alcohol dependence were available from the Global Information System on Alcohol and Health (World Health Organization, 2019), the sensitivity and specificity of the AUDIT as a screening instrument for AUDs, inclusively, and alcohol dependence, independent of other AUDs, were estimated for all AUDIT cut-points (i.e., 0 to 40), by sex. For those countries where no standard drink size was found through the systematic reviews (Kalinowski and Humphreys, 2016; Mongan and Long, 2015), the standard drink size was assumed to be 10 g of EtOH (based on the global standard drink size; World Health Organization, 2018a). Data on sensitivity and specificity were then combined with data on the prevalence of AUDs, inclusively, and alcohol dependence, independent of other AUDs, in the general population (as obtained from the Global Information System on Alcohol and Health) for 2016, the most

recent year for which such data are available (World Health Organization, 2019). Information on the procedures used to estimate the prevalence of AUD is available on the Global Information System on Alcohol and Health (World Health Organization, 2019). These data were used to estimate the number of true-positive, false-positive, true-negative, and false-negative proportions, where the term positive refers to having an AUDIT score above a specified cut-point and true positive refers to both having an AUDIT score above a specified cut-point and having an AUD. These data were also used to estimate the number of people needed to be screened in order to treat 1 individual with an AUD, inclusively, or with alcohol dependence, independent of other AUDs. The 95% uncertainty interval (UI) for each respective point estimate was constructed using a Monte Carlo–like algorithm, generating 1,000 simulations which were estimated by sampling the lowest level parameters from their respective error distributions, and using the 2.5th and 97.5th percentiles of the resulting distribution as the UI. All statistical analyses were performed using R version 3.4.1 (R Core Team, 2017).

RESULTS

The electronic search yielded a total of 967 articles, and 24 articles were identified through the manual search. After removing 540 duplicate articles, 451 articles were screened using titles and abstracts. One hundred and eighty full-text articles were retrieved for further consideration, 144 of which were subsequently excluded. Thus, a total of 36 studies were retained for data extraction. The 36 studies contained data for a total of 20 countries: Argentina ($n = 1$), Belgium ($n = 2$), Brazil ($n = 1$), China ($n = 1$), France ($n = 2$), Germany ($n = 1$), India ($n = 1$), Mexico ($n = 1$), Nepal ($n = 1$), Netherlands ($n = 1$), Norway ($n = 1$), Poland ($n = 1$), South Korea ($n = 4$), Spain ($n = 2$), Sri Lanka ($n = 1$), Sweden ($n = 2$), Switzerland ($n = 1$), Turkey ($n = 1$), the United Kingdom (UK; $n = 2$), and the United States ($n = 10$; 1 study reported on a sample drawn from France and Switzerland). Thirty studies reported the classification accuracy of the AUDIT with respect to AUDs, and 20 studies reported the classification accuracy of the AUDIT with respect to alcohol dependence. The characteristics of each included study are presented in Table S1 (men) and Table S2 (women). Of the 36 studies included, the majority were considered to be high risk for bias and low risk for applicability (Table S3). A schematic diagram of the search strategy is depicted in Fig. 1.

With respect to the included studies, Fig. 2 depicts the relationship between the AUDIT cut-point and the sensitivity and specificity of the AUDIT in identifying AUDs, by sex, assuming a standard drink size of 10 g of EtOH (see Fig. S1 for alcohol dependence). The heterogeneity estimates for the sensitivity and specificity of the AUDIT in identifying AUDs and alcohol dependence are presented in Table S4. As depicted in the funnel plots for the sensitivity and specificity estimates of the AUDIT in identifying AUDs (Fig. S2) and alcohol dependence (Fig. S3), there was no indication of publication bias.

It should be noted that the description of the findings presented below focuses on the sex-specific screening statistics for AUDs using the AUDIT cut-point score which provided

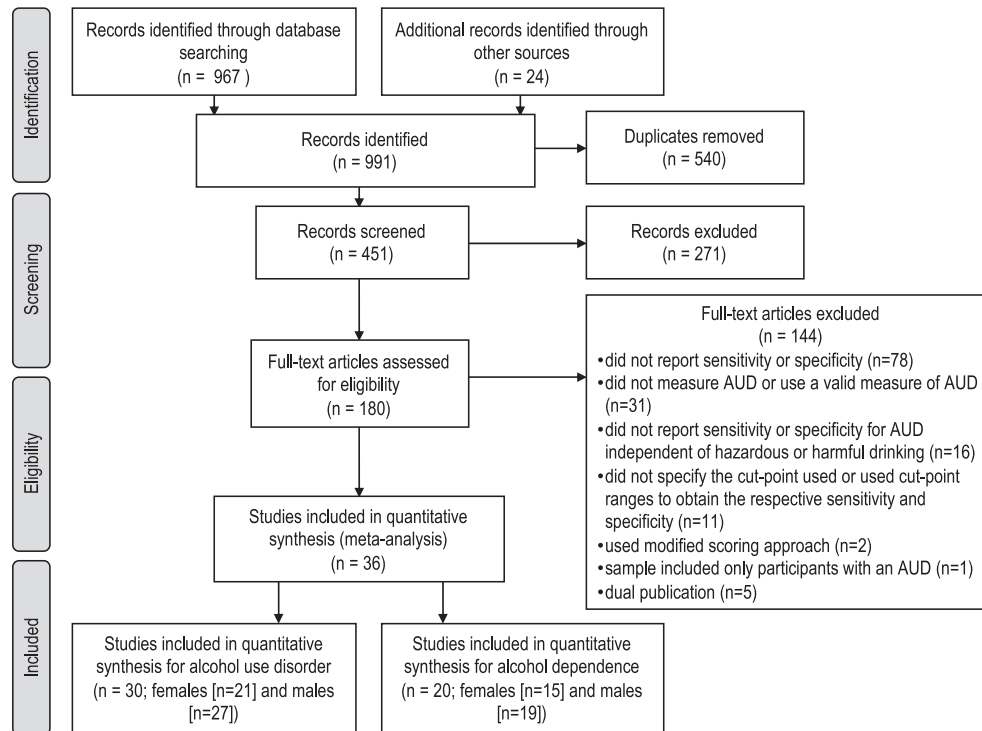


Fig. 1. Schematic diagram of the search strategy.

a specificity of 0.95 for the largest country in each Institute for Health Metrics and Evaluation (IHME) region and in the UK. However, the sex-specific screening statistics for both AUDs and alcohol dependence using the AUDIT cut-points which provided a specificity of 0.90, 0.95, and 0.98 are presented for all 190 countries in Table S5 (AUDs) and Table S6 (alcohol dependence).

The Use of the AUDIT in Identifying AUDs Among Men

At a specificity of 0.95 on the AUDIT, among the largest countries in each IHME region and the UK, the 3 countries that had the highest true-positive rates among men were Russia (299 per 1,000 [95% UI: 267 to 332 per 1,000]), Poland (184 per 1,000 [95% UI: 158 to 208 per 1,000]), and the United States (132 per 1,000 [95% UI: 112 to 153 per 1,000]), which was mirrored by having had the lowest number of men needed to be screened in order to treat 1 man with an AUD: 3 (95% UI: 3 to 4), 5 (95% UI: 5 to 6), and 8 (95% UI: 7 to 9), respectively (Fig. 3). All 3 of these countries had a higher relative prevalence of AUDs among men (compared to other countries). Likewise, the 3 countries with a relatively low prevalence of AUDs among men had the lowest true-positive rates and highest number of men needed to be screened in order to treat 1 man with an AUD: Nigeria (9 per 1,000 [5 to 16 per 1,000] and 116 [95% UI: 61 to 201], respectively), Indonesia (11 per 1,000 [6 to 19 per 1,000] and 91 [95% UI: 51 to 155], respectively), and Egypt (21 per 1,000 [14 to 30 per 1,000] and 48 [95% UI: 34 to 71], respectively).

As depicted in Fig. 3, the larger the standard drink size used, the higher the AUDIT cut-point that was needed to produce a specificity of 0.95. For instance, when using a standard drink size of 8 g of EtOH (UK), an AUDIT cut-point of 17 was needed to produce a specificity of 0.95, while when using a standard drink size of 14 g of EtOH (Mexico and the United States), an AUDIT cut-point of 14 was needed to produce a specificity of 0.95. This relationship was reflected in the beta coefficients in the regression (Table S7), where standard drink size was found to significantly affect the sensitivity ($\beta = -0.021$) and specificity ($\beta = 0.016$) of the AUDIT in detecting AUDs among men. However, this relationship was not found for alcohol dependence, where an AUDIT cut-point of 18 was needed to produce a specificity of 0.95 in all countries, regardless of standard drink size (Fig. S4); again, this relationship was reflected in the regression, with nonsignificant beta coefficients for both sensitivity and specificity in detecting alcohol dependence among men (Table S7).

The Use of the AUDIT in Identifying AUDs Among Women

At a specificity of 0.95, the AUDIT performed poorly in accurately identifying women with an AUD, as indicated by the low true-positive rate relative to the high false-positive rate (Fig. 4). This was likely due to the low prevalence of AUDs among women (compared to men), as the false-positive rate is dependent on prevalence (i.e., false-positive rate = $1 - \text{specificity} * 1 - \text{prevalence}$). These findings assume the correct application and scoring of the AUDIT and the

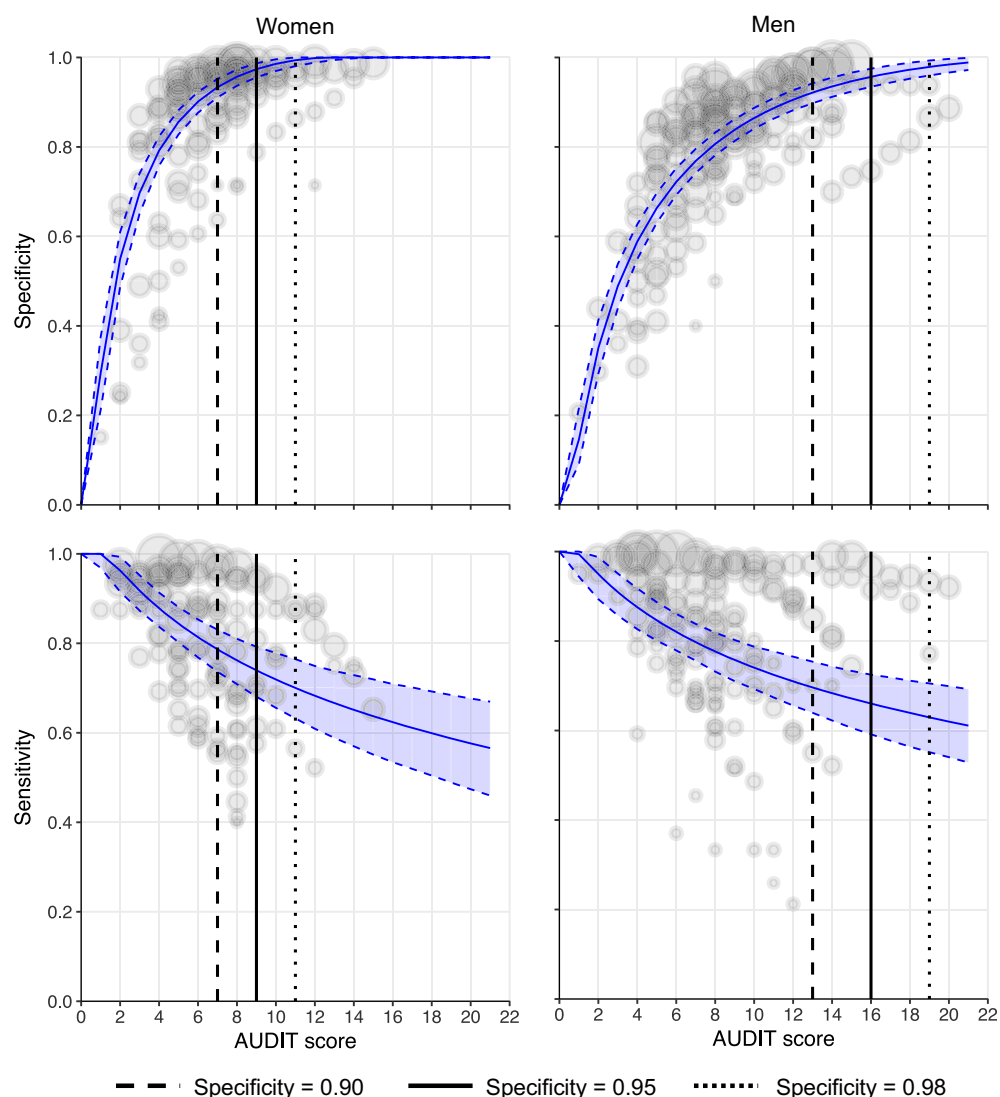


Fig. 2. Relationship between the AUDIT cut-point and the sensitivity and specificity of the AUDIT in identifying AUDs, by sex, assuming a standard drink size of 10 g of EtOH. *Note.* Audit cut-points were rounded to the nearest integer. Specificity = 0.90; women (AUDIT score = 7; sensitivity = 0.848); men (AUDIT score = 13; sensitivity = 0.829). Specificity = 0.95; women (AUDIT score = 9; sensitivity = 0.810); men (AUDIT score = 16; sensitivity = 0.810). Specificity = 0.96; women (AUDIT score = 11; sensitivity = 0.773); men (AUDIT score = 19; sensitivity = 0.792).

correct application of AUD diagnostic criteria. With respect to the largest countries in each IHME region and the UK, the only 2 countries where the true-positive rates were higher than the false-positive rates were Russia and the United States—2 countries with a relatively higher prevalence of AUDs among women (compared to other countries). Further, the number of women needed to be screened in order to treat 1 woman with an AUD ranged from 14 (95% UI: 11 to 17) in the United States to 1,025 (95% UI: 182 to 6,222) in Nigeria—again, reflective of the prevalence of AUDs in the respective countries. The poor performance of the AUDIT in identifying women with an AUD was likely also due to the low cut-point (9) that provided a specificity of 0.95, suggesting that the current cut-points used to identify women most in need of AUD treatment are not effective in accurately doing so and, in fact, are resulting in high false-positive rates.

With respect to standard drink size, the performance of the AUDIT in accurately identifying women with an AUD was not impacted by the standard drink size used in a country, as indicated by a cut-point of 9 providing a specificity of 0.95 in all countries irrespective of their standard drink size (Fig. 4). This was also reflected in the beta coefficients in the regression (Table S7), where standard drink size was found to significantly affect the sensitivity ($\beta = -0.030$) of the AUDIT in detecting AUDs among women, but not the specificity. However, with respect to alcohol dependence, the performance of the AUDIT in accurately identifying women with alcohol dependence was impacted by the standard drink size (Fig. S5). This was, again, reflected in the beta coefficients in the regression, where standard drink size was found to significantly affect the sensitivity ($\beta = -0.077$) and

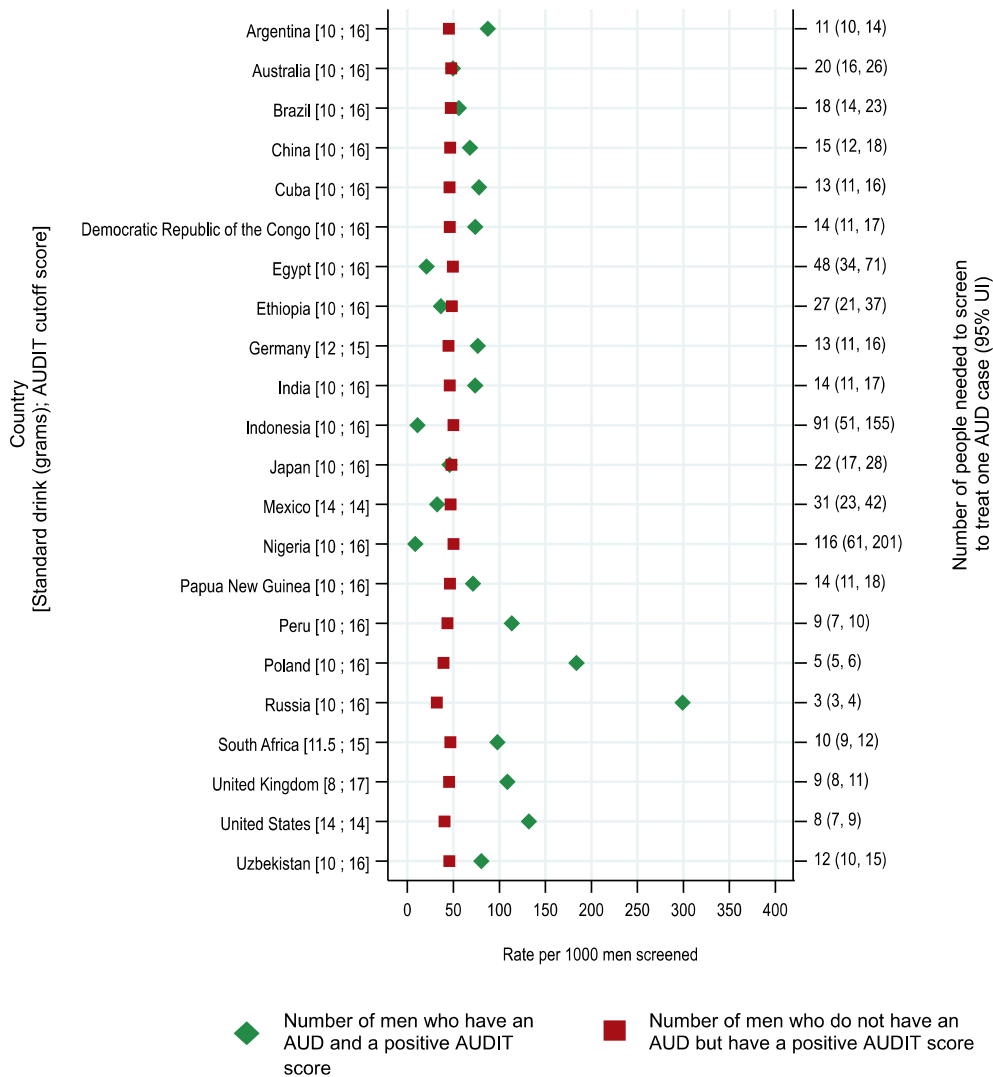


Fig. 3. Screening statistics, by country, for AUDs per 1,000 men screened using the AUDIT cut-point that provides a specificity of 0.95. *Note.* Rates are based on 1,000 men screened and do not represent strata of men with and without AUDs.

specificity ($\beta = 0.036$) of the AUDIT in detecting alcohol dependence among women (Table S7).

DISCUSSION

This is the first study to quantitatively pool all available data on the classification accuracy of the AUDIT with respect to its ability to screen individuals for an AUD, inclusively, as well as its ability to screen individuals for alcohol dependence (the most severe AUD), and to provide country- and sex-specific screening statistics based on the standard drink size of each respective country. The accuracy of screening tools is affected by factors such as demographics (sex), point of diagnosis (current vs. lifetime), and the criterion standard to which it is compared—all of which were accounted for in the current study.

We found that in countries with a low prevalence of AUDs, the number of people needed to be screened using the

AUDIT in order to treat 1 individual with an AUD was very high. This finding calls into question whether the AUDIT should be recommended for use as a screening tool for AUDs in such countries, especially given the scarce resources for health care, and at a time when primary health care is already overburdened. For instance, in China, the most populous country in the world, the number of people needed to be screened in order to treat 1 individual with an AUD would be 15 for men and 618 for women. If the expectation is that patients with an AUD will be treated by their primary health care provider, an oral AUDIT interview is recommended (Babor et al., 2001). Assuming the AUDIT takes 10 minutes to administer orally in person, including a proper demonstration of standard drink sizes, this would require a health care worker to spend 2.5 hours in order to identify 1 man in need of AUD treatment, and 103 hours in order to identify 1 woman in need of AUD treatment. These numbers do not indicate a high efficiency in use, or a cost-effective use

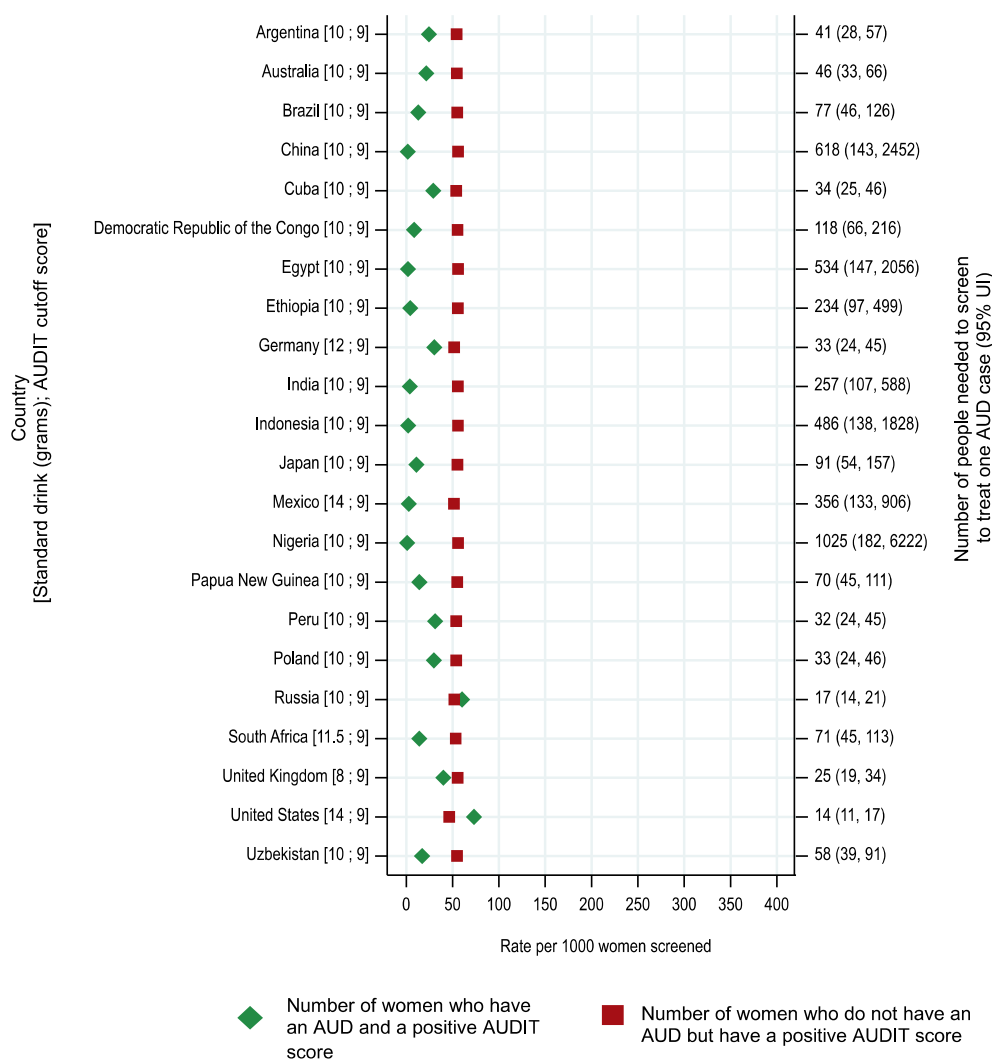


Fig. 4. Screening statistics, by country, for AUDs per 1,000 women screened using the AUDIT cut-point that provides a specificity of 0.95. *Note.* Rates are based on 1,000 women screened and do not represent strata of women with and without AUDs.

of staff time. While it may be possible for the AUDIT to be administered via a different mode (telephone, Web-based, or self-administration; DiGiuseppi et al., 2006; Graham et al., 2007; Kypri et al., 2008) prior to meeting with a primary health care professional, an adaptation of the AUDIT (e.g., the AUDIT-C [Bush et al., 1998]), or even a simple, single-item screening question about alcohol use, may, in fact, be more appropriate for screening purposes (Smith et al., 2009). Regardless, a question-based screening approach would need to ensure patient understanding and be sensitive to gender and cultural differences. In addition, the continued use of the AUDIT to screen for AUDs, especially among women, is likely to result in unnecessary stigmatization given the false-positive rates. AUDs are highly stigmatized, and while progress has been shown for other mental disorders, the stigmatization of AUDs has remained high over the past decades (Schomerus et al., 2011, 2014).

The different definitions of AUDs present additional problems, and in fact, the current findings may be due, in part, to

deficiencies in the DSM and ICD diagnostic criteria for AUDs. In their most recent revisions, the DSM and ICD have drifted further apart in their definitions of AUDs, and the correlation between both systems is not sufficient (Degenhardt et al., 2019; Room and Rehm, 2019). In fact, the DSM-5 has been questioned for its utility in the clinical practice of AUD management (Carvalho et al., 2019). As a means of classifying patients with respect to severity, prognosis, and appropriate treatment(s), these systems are falling short when it comes to AUDs, as the behaviors used to define AUDs are not only vaguely described but the behaviors themselves have changed over time. Additional attention to developing uniform diagnostic systems, as well as ensuring that the criteria used in standard tests are reflective of the respective diagnostic systems, will ultimately help to advance screening efforts.

Indeed, the current situation is that the overwhelming majority of individuals with an AUD are not receiving any treatment (Blanco et al., 2015; Kohn et al., 2004; Rehm

et al., 2015). Rather than using a screening tool, like the AUDIT, that requires too many resources, another strategy would be to screen only those individuals with common diseases and/or harms highly associated with heavy alcohol use and AUDs (for an overview of alcohol use and disease, see Rehm et al., 2017b). At the primary health care level, hypertension (Rehm et al., 2017a), insomnia, and injuries (Israel et al., 1996) are frequently observed examples of the consequences of heavy alcohol use and AUDs, and if screening for AUDs were to be restricted to individuals presenting with these and other commonly associated diseases and/or harms, screening would be more focused and might result in higher true-positive rates (individuals screening positive who need intervention). It should be noted that all of the above considerations relate to AUDs specifically, and not to hazardous or harmful alcohol use in general.

Given the recent attention that the AUDIT has received in relation to phenotyping individuals with AUDs, as well as hazardous and harmful alcohol use, in genome-wide association studies (Justice et al., 2018; Sanchez-Roige et al., 2019), the current study is timely. In fact, there is some indication that the AUDIT could be used to identify individuals genetically at risk for AUDs.

Limitations

There are a few limitations that should be acknowledged. The included studies used 1 or the other of the diagnostic systems (ICD or DSM), which, as discussed above, lack concordance with respect to AUDs. This could have introduced an unknown level of heterogeneity into the pooled estimates. Further, the vast majority of included studies did not report the standard drink size used, and there was no indication of any pilot testing prior to the implementation of the AUDIT to determine whether the concept of a standard drink size was understandable to study participants. The findings of the current study are based on country-specific drink sizes, which may or may not be reflective of those used in the individual studies. Lastly, there was inconsistent adherence among the included studies to the methodological standards designed to increase the validity of studies on diagnostic test performance, and it was simply not possible to restrict the analysis to those studies that were considered to be low risk for bias. This lack of adherence is reflective of the difficulty of conducting research in a clinical setting; however, future studies should be mindful of these standards.

Utility of an Alcohol Use Disorder Diagnosis

It must be noted that our study focused on AUDs and did not include studies that evaluated the performance of the AUDIT in identifying hazardous and harmful alcohol use. The AUDIT is a 10-item questionnaire which was developed by the WHO in 1989 as a screening tool specifically designed to identify individuals who were most in need of: (i) simple

advice focused on reducing hazardous drinking, (ii) brief counseling and continued monitoring, and/or (iii) further diagnostic evaluations and referrals to specialized treatment (Babor et al., 2001). Aside from being used to identify individuals who have a hazardous (or risky) pattern of drinking, as per its original purpose, the AUDIT is often used, in both clinical and research settings, to ascertain whether individuals are likely to meet the diagnostic criteria for an AUD.

The term AUD is defined either by the DSM, currently in its fifth edition (American Psychiatric Association, 2013), which uses a single diagnostic term (AUD) with a severity indicator, or by the ICD, currently in its 11th revision (World Health Organization, 2016), which has retained a division into 2 main diagnoses (harmful use of alcohol and alcohol dependence syndrome) to indicate an AUD. Under the DSM-5, anyone meeting any 2 of the following 11 criteria within a 12-month period would receive a diagnosis of having an AUD (with severity—mild, moderate, or severe—being based on the number of criteria met): Alcohol is consumed in larger amounts over a longer period than intended; persistent desire or unsuccessful attempts to cut down or control one's alcohol use; large amount of time spent in activities necessary to obtain, use, or recover from the effects of alcohol; cravings; failure to fulfill major obligations at work, school, or home; continued use despite persistent or recurrent social or interpersonal problems; social, occupational, or recreational activities given up or reduced; use in situations in which it is physically hazardous; use continued despite persistent or recurrent physical or psychological problems; increased tolerance; and withdrawal symptoms. With respect to the ICD, the harmful use of alcohol refers to alcohol consumption that results in physical and mental health consequences, while alcohol dependence is a maladaptive pattern of drinking that leads to clinically significant impairment or distress, such as a strong desire to consume alcohol, impaired control over its use, persistent drinking despite harmful consequences, a higher priority given to drinking than to other social, occupational, or recreational activities and obligations, increased alcohol tolerance, and a physical withdrawal reaction when alcohol use is discontinued (known as alcohol withdrawal syndrome).

Thus, our finding that the AUDIT does not perform well as a universal screening tool for identifying individuals with an AUD cannot be extended to include hazardous and harmful alcohol use. Accordingly, the discussion above is specific to AUDs and does not necessarily apply to hazardous and harmful patterns of alcohol use. This pattern of alcohol consumption can be extremely detrimental in and of itself (Rehm et al., 2017b), and a rigorous quantitative analysis of the AUDIT's ability to identify such hazardous and harmful patterns of use should be conducted in the future. Further, the current study was limited to the performance of the full AUDIT; however, it is possible that certain items of the AUDIT (e.g., consumption items or problem items) would prove to be clinically useful in screening for AUDs, and this possibility warrants further evaluation.

CONCLUSION

In conclusion, we do not recommend the AUDIT as a universal screening tool for identifying individuals with an AUD, especially in countries and among populations with a low AUD prevalence (e.g., among women). There are a number of factors that may contribute to the poor performance of the AUDIT in the screening of individuals for an AUD, such as cultural and social ideologies, as well as certain items inordinately influencing the overall psychometric characteristics. Regardless, different forms of screening for AUDs should be considered and evaluated for use in primary care. A more targeted approach to screening based on the presence of other medical conditions would also reduce the stigmatization of AUDs, as screening, advice, and treatment would be seen as part of the routine management of conditions customarily dealt with at this level of health care.

FUNDING

This research was financially supported by the Pan American Health Organization, and the Canadian Institutes of Health Research, Institute of Neurosciences, Mental Health and Addiction (CRISM Ontario Node grant number SMN-13950).

CONFLICT OF INTEREST

None to declare.

REFERENCES

- American Psychiatric Association (2013) Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition: DSM-5. Author, Arlington, VA.
- Babor TF, Higgins-Biddle JC, Saunders JB, Monteiro MG (2001) The Alcohol Use Disorders Identification Test: Guidelines for Use in Primary Care. World Health Organization, Department of Mental Health and Substance Dependence, Geneva, Switzerland.
- Babor TF, Robaina K (2016) The Alcohol Use Disorders Identification Test (AUDIT): A review of graded severity algorithms and national adaptations. *Int J Alcohol Drug Res* 5:17–24.
- Blanco C, Iza M, Rodriguez-Fernandez JM, Baca-Garcia E, Wang S, Olfson M (2015) Probability and predictors of treatment-seeking for substance use disorders in the U.S. *Drug Alcohol Depend* 149:136–144.
- Bush K, Kivlahan DR, McDonnell MB, Fihn SD, Bradley KA (1998) The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. *Arch Intern Med* 158:1789–1795.
- Carvalho A, Heilig M, Perez A, Probst C, Rehm J (2019) Alcohol use disorders. *Lancet* 394.
- Deeks JJ, Higgins JPT, Altman DG (2008) Identifying and assessing heterogeneity, in *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins JPT, Green S eds). John Wiley & Sons, Chichester, UK.
- Degenhardt L, Bharat C, Bruno R, Glantz MD, Sampson NA, Lago L, Aguilar-Gaxiola S, Alonso J, Andrade LH, Bunting B, Caldas-de-Almeida JM, Cia AH, Gureje O, Karam EG, Khalaf M, McGrath JJ, Moskalewicz J, Lee S, Mneimneh Z, Navarro-Mateu F, Sasu CC, Scott K, Torres Y, Poznyak V, Chatterji S, Kessler RC; WHO World Mental Health Survey Collaborators (2019) Concordance between the diagnostic guidelines for alcohol and cannabis use disorders in the draft ICD-11 and other classification systems: analysis of data from the WHO's World Mental Health Surveys. *Addiction* 114:534–552.
- DerSimonian R, Laird N (1986) Meta-analysis in clinical trials. *Control Clin Trials* 7:177–188.
- Deville WL, Buntinx F, Bouter LM, Montori VM, de Vet HC, van der Windt DA, Bezemer PD (2002) Conducting systematic reviews of diagnostic studies: didactic guidelines. *BMC Med Res Methodol* 2:9.
- DiGiuseppi C, Goss C, Xu S, Magid D, Graham A (2006) Telephone screening for hazardous drinking among injured patients seen in acute care clinics: feasibility study. *Alcohol Alcohol* 41:438–445.
- Egger M, Smith GD, Schneider M, Minder C (1997) Bias in meta-analysis detected by a simple, graphical test. *BMJ* 315:629–634.
- Freeman MF, Tukey JW (1950) Transformations related to the angular and the square root. *Ann Math Stat* 21:607–611.
- Graham A, Goss C, Xu S, Magid DJ, DiGiuseppi C (2007) Effect of using different modes to administer the AUDIT-C on identification of hazardous drinking and acquiescence to trial participation among injured patients. *Alcohol Alcohol* 42:423–429.
- Hasin D, Hatzenbuehler ML, Keyes K, Ogburn E (2006) Substance use disorders: Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV) and International Classification of Diseases, tenth edition (ICD-10). *Addiction* 101(Suppl. 1):59–75.
- Higgins JP, Thompson SG (2002) Quantifying heterogeneity in a meta-analysis. *Stat Med* 21:1539–1558.
- Israel Y, Hollander O, Sanchez-Craig M, Booker S, Miller V, Gingrich R, Rankin JG (1996) Screening for problem drinking and counseling by the primary care physician-nurse team. *Alcohol Clin Exp Res* 20:1443–1450.
- Justice AC, Smith RV, Tate JP, McGinnis K, Xu K, Becker WC, Lee KY, Lynch K, Sun N, Concato J, Fiellin DA, Zhao H, Gelernter J, Kranzler HR (2018) AUDIT-C and ICD codes as phenotypes for harmful alcohol use: association with ADH1B polymorphisms in two US populations. *Addiction* 113:2214–2224.
- Kalinowski A, Humphreys K (2016) Governmental standard drink definitions and low-risk alcohol consumption guidelines in 37 countries. *Addiction* 111:1293–1298.
- Kohn R, Saxena S, Levav I, Saraceno B (2004) The treatment gap in mental health care. *Bull World Health Organ* 82:858–866.
- Kypri K, Langley JD, Saunders JB, Cashell-Smith ML, Herbison P (2008) Randomized controlled trial of web-based alcohol screening and brief intervention in primary care. *Arch Intern Med* 168:530–536.
- Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J, Moher D (2009) The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med* 6:e1000100.
- Manthey J, Shield K, Rylett M, Hasan OSM, Probst C, Rehm J (2019) Alcohol exposure between 1990 and 2017 and forecasts until 2030: a global modelling study. *Lancet* 393:2493–2502.
- Mongan D, Long J (2015) Standard Drink Measures in Europe. Peoples' Understanding of Standard Drinks and Their Use in Drinking Guidelines, Alcohol Surveys and Labelling. Health Research Board, Dublin, Ireland.
- National Institute on Alcohol Abuse and Alcoholism (2003) *Assessing Alcohol Problems: A Guide for Clinicians and Researchers*, 2nd Edition. NIH Publication No. 03–3745. Author, Bethesda, MD.
- R Core Team (2017) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Rehm J, Anderson P, Manthey J, Shield K, Struzzo P, Wojnar M, Gual A (2016) Alcohol use disorders in primary health care: what do we know and where do we go? *Alcohol Alcohol* 51:422–427.
- Rehm J, Anderson P, Prieto JAA, Armstrong I, Aubin HJ, Bachmann M, Bastus NB, Brotons C, Burton R, Cardoso M, Colom J, Duprez D, Gmel G, Gual A, Kraus L, Kreutz R, Liira H, Manthey J, Møller L, Okruhlica L, Roerecke M, Scafato E, Schulte B, Segura-Garcia L, Shield KD, Sierra C, Vyshinskiy K, Wojnar M, Zarco J (2017a) Towards new recommendations to reduce the burden of alcohol-induced hypertension in the European Union. *BMC Med* 15:173.
- Rehm J, Gmel GE Sr, Gmel G, Hasan OSM, Imtiaz S, Popova S, Probst C, Roerecke M, Room R, Samokhvalov AV, Shield K, Shuper PA (2017b)

- The relationship between different dimensions of alcohol use and the burden of disease—an update. *Addiction* 112:968–1001.
- Rehm J, Manthey J, Struzzo P, Gual A, Wojnar M (2015) Who receives treatment for alcohol use disorders in the European Union? A cross-sectional representative study in primary and specialized health care. *Eur Psychiatry* 30:885–893.
- Rehm J, Room R, Graham K, Monteiro M, Gmel G, Sempos CT (2003) The relationship of average volume of alcohol consumption and patterns of drinking to burden of disease: an overview. *Addiction* 98:1209–1228.
- Rehm J, Shield KD, Gmel G, Rehm MX, Frick U (2013) Modeling the impact of alcohol dependence on mortality burden and the effect of available treatment interventions in the European Union. *Eur Neuropsychopharmacol* 23:89–97.
- Room R, Rehm J (2019) Commentary on Degenhardt et al. (2019): harm to others matters in substance use disorders, and so does discordance between the diagnostic systems. *Addiction* 114:553–554.
- Sanchez-Roige S, Palmer AA, Fontanillas P, Elson SL, Adams MJ, Howard DM, Edenberg HJ, Davies G, Crist RC, Deary IJ, McIntosh AM, Clarke TK (2019) Genome-wide association study meta-analysis of the Alcohol Use Disorders Identification Test (AUDIT) in two population-based cohorts. *Am J Psychiatry* 176:107–118.
- Saunders JB, Aasland OG, 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.
- Schomerus G, Lucht M, Holzinger A, Matschinger H, Carta MG, Angermeyer MC (2011) The stigma of alcohol dependence compared with other mental disorders: a review of population studies. *Alcohol Alcohol* 46:105–112.
- Schomerus G, Matschinger H, Angermeyer MC (2014) Attitudes towards alcohol dependence and affected individuals: persistence of negative stereotypes and illness beliefs between 1990 and 2011. *Eur Addict Res* 20:293–299.
- Smith PC, Schmidt SM, Allensworth-Davies D, Saitz R (2009) Primary care validation of a single-question alcohol screening test. *J Gen Intern Med* 24:783–788.
- Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, Becker BJ, Sipe TA, Thacker SB (2000) Meta-analysis of observational studies in epidemiology: a proposal for reporting. *J Am Med Assoc* 283:2008–2012.
- Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM, QUADAS-2 Group (2011) QUADAS-2: a revised tool for the Quality Assessment of Diagnostic Accuracy Studies. *Ann Intern Med* 155:529–536.
- World Health Organization (2016) ICD-10 Version: 2016. Available at <https://icd.who.int/browse10/2016/en>. Accessed December 12, 2018.
- World Health Organization (2018a) WHO Global Coordination Mechanism on the Prevention and Control of NCDs: NCD and the Sustainable Development Goals. Available at <http://www.who.int/global-coordination-mechanism/ncd-themes/sustainable-development-goals/en/>. Accessed December 6, 2018.
- World Health Organization (2018b) Global Status Report on Alcohol and Health 2018. World Health Organization, Geneva, Switzerland.
- World Health Organization (2018c) Management of Substance Abuse: WHO Launches SAFER Alcohol Control Initiative to Prevent and Reduce Alcohol-Related Death and Disability: WHO-Led Initiative and Action Package Aim to Support Global Target of Reducing Harmful Use of Alcohol by 10% by 2025 [press release]. WHO Department of Communications, New York.
- World Health Organization (2019) Global Information System on Alcohol and Health. World Health Organization, Geneva, Switzerland.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Identified studies that validated the AUDIT for identifying alcohol use disorders and alcohol dependence among males.

Table S2. Identified studies that validated the AUDIT for identifying alcohol use disorders and alcohol dependence among females.

Table S3. Risk of bias and applicability assessments of the identified studies that validated the AUDIT for identifying alcohol use disorders and alcohol dependence.

Table S4. Heterogeneity estimates for the sensitivity and specificity of the Alcohol Use Disorder Identification Test in identifying alcohol use disorders and alcohol dependence.

Table S5. Screening statistics for alcohol use disorders using an AUDIT cut-point which provides specificities of 0.90, 0.95, and 0.98 by country and sex.

Table S6. Screening statistics for alcohol dependence using an AUDIT cut-point which provides specificities of 0.90, 0.95, and 0.98 by country and sex.

Table S7. Statistical models for estimating the sensitivity and specificity of the AUDIT as a screening instrument for alcohol use disorders and alcohol dependence.

Fig. S1. Relationship between the AUDIT cut-point and the sensitivity and specificity of the AUDIT in identifying alcohol dependence by sex, assuming a standard drink size of 10 grams.

Fig. S2. Funnel plots for Freeman-Tukey Double arcsine transformed sensitivity and specificity estimates for the AUDIT in identifying alcohol use disorders.

Fig. S3. Funnel plots for Freeman-Tukey Double arcsine transformed sensitivity and specificity estimates for the AUDIT in identifying alcohol dependence.

Fig. S4. Screening statistics by country for alcohol dependence per 1,000 men screened using the AUDIT cut-point that provides a specificity of 0.95.

Fig. S5. Screening statistics by country for alcohol dependence per 1,000 women screened using the AUDIT cut-point that provides a specificity of 0.95.