Annals of Internal Medicine | Review

Meta-analysis: Are 3 Questions Enough to Detect Unhealthy Alcohol Use?

Levente Kriston, MA; Lars Hölzel, MA; Ann-Kristin Weiser, BA; Michael M. Berner, MD; and Martin Härter, MD, PhD

Background: Both the 10-item Alcohol Use Disorders Identification Test (AUDIT) and its abbreviated 3-item version (Alcohol Use Disorders Identification Test–Consumption [AUDIT-C]) are considered to detect unhealthy alcohol use accurately.

Purpose: To examine whether the AUDIT-C is as accurate as the full AUDIT for detecting unhealthy alcohol use in adults.

Data Sources: MEDLINE, EMBASE, CINAHL, Web of Science, PsycINFO, and BIOSIS Previews from 1998 to July 2008.

Study Selection: Three independent reviewers selected studies that administered both the AUDIT and the AUDIT-C, applied a valid reference standard, avoided verification and incorporation bias, and reported relevant data. No language restrictions were applied.

Data Extraction: Two reviewers extracted study characteristics and outcome data, which were cross-checked by a third reviewer. One reviewer assessed methodological quality with a standardized checklist.

Data Synthesis: Fourteen studies were found. Most involved primary care patients in Europe and the United States. Sample sizes ranged between 112 and 13 438 patients, and sex and age distributions varied considerably. No statistically significant differences were found between the overall accuracy of the AUDIT and the AUDIT-C for detecting risky drinking, alcohol use disorders, or unhealthy alcohol use in primary care. Hierarchical summary receiver-operating characteristic curve analysis yielded pooled positive likelihood ratios of 6.62 for the AUDIT and 2.99 for the AUDIT-C, respectively, for detecting risky drinking; 4.03 and 3.82, respectively, for detecting any alcohol use disorder; and 4.82 and 3.91, respectively, for detecting risky drinking or any alcohol use disorder. Findings from a few studies on general population samples and inpatients suggested but did not prove that the AUDIT might be better than the AUDIT-C for identifying severe conditions, such as alcohol dependence.

Limitation: Studies used different reference standards and had heterogeneous findings.

Conclusion: Available evidence is inconclusive but suggests that the full AUDIT may be superior to the AUDIT-C for identifying unhealthy alcohol use in adults in some settings.

Ann Intern Med. 2008;149:879-888. www.annals.org
For author affiliations, see end of text.

Alcohol use disorders, such as abuse and dependence, have serious medical consequences. One in 12 persons in the United States is affected, with the 12-month prevalence of alcohol dependence at 3.8% and that of alcohol abuse at 4.7%. A history of alcohol dependence is present in 12.5% of the population, and the lifetime prevalence of alcohol abuse is 17.8% (1). The prevalence of risky drinking (unhealthy alcohol use without an alcohol use disorder) in outpatients ranges from 7% to 29% for men and from 3% to 10% in women (2). In general medical inpatients and emergency departments, these rates are even higher, reaching up to 50% (3–7). The annual economic costs to society associated with unhealthy alcohol use are estimated to be approximately $150 billion (7–10).

To increase recognition of patients with unhealthy alcohol use by health care professionals, leading health care organizations and experts propose the administration of screening questionnaires (7, 11–15). One of the most frequently recommended instruments is the Alcohol Use Disorders Identification Test (AUDIT) from the World Health Organization (16). The full AUDIT consists of 10 items and has been extensively researched in several settings and countries (17). The AUDIT has proved to be a psychometrically sound instrument for screening for all forms of unhealthy alcohol use (risky drinking, alcohol abuse, and alcohol dependence) (17, 18).

In 1998, a promising study was published (19) of an abbreviated version of the AUDIT questionnaire that consisted of the first 3 items only. Because these items are related to alcohol intake, the brief instrument is called the Alcohol Use Disorders Identification Test–Consumption (AUDIT-C). Several studies concluded that this 3-item version is also appropriate to screen for unhealthy alcohol use (17). However, although some studies (19–21) suggest statistically significant differences between the screening properties of the AUDIT-C and the AUDIT in certain settings, this difference has not undergone systematic examination. This topic is of special relevance because administration of the AUDIT-C instead of the AUDIT could save substantial health care resources. We therefore sought to examine whether the AUDIT-C is as accurate as the AUDIT for detecting unhealthy alcohol use in adults.

See also:

Print
Editors' Notes 880
Editorial comment. 904

Web-Only
Appendix Tables
Appendix Figures
CME quiz
Conversion of graphics into slides

Context

Does a 3- or a 10-item questionnaire (Alcohol Use Disorders Identification Test–Consumption [AUDIT-C] or Alcohol Use Disorders Identification Test [AUDIT]) better identify unhealthy alcohol use in adults?

Contribution

This meta-analysis found 14 studies that directly compared the accuracy of the AUDIT-C with that of the AUDIT. The studies, conducted mainly in primary care settings, used different reference standards and had heterogeneous findings. Some pooled analyses suggested that primary care patients with a positive AUDIT screening result were more likely than those with a positive AUDIT-C screening result to be engaging in risky drinking.

Implication

This body of evidence suggests, but does not prove, that the AUDIT may be better than the AUDIT-C in some settings for identifying unhealthy alcohol use in adults.

—The Editors

METHODS

Findings are reported according to the QUOROM (Quality of Reporting of Meta-Analysis) statement (22).

Data Sources and Searches

We searched MEDLINE, EMBASE, CINAHL, Web of Science, PsycINFO, and BIOSIS Previews databases by using the terms "AUDIT-C" and "AUDIT-3" with no language restrictions from 1998 to July 2008. In addition, we retrieved all studies registered in Web of Science that cited the original publication on the AUDIT-C (19) until July 2008. We checked reference lists of the included studies and reviews on AUDIT (14, 17, 18, 23) for additional relevant publications.

Study Selection

Two reviewers screened the title and abstract of the retrieved references to identify possibly relevant publications. Subsequently, 3 independent reviewers checked the full text of all potentially eligible studies to decide about inclusion. The following inclusion criteria had to be fulfilled: diagnostic accuracy studies that administered both the AUDIT and the AUDIT-C, use of a valid reference standard of any unhealthy alcohol use, receipt of verification by using a reference standard by the whole sample or a random selection of it (avoidance of partial verification), receipt of the same reference standard regardless of the index test result (avoidance of differential verification), independence of index test and reference standard (avoidance of incorporation bias), and reporting of sufficient data to reproduce the diagnostic 2 × 2 contingency table (number of true-positive, true-negative, false-positive, and false-negative cases) for both tests. We applied no restrictions on language, population, or setting.

To provide an exact description of the reference standards, we used standardized terminology (7). We examined 4 possible reference standards that described a current or active condition (**Table 1**): risky drinking, any alcohol use disorder, alcohol dependence, and any unhealthy alcohol use. To date, no uniform terminology has been established for risky drinking. It is frequently termed as heavy, excessive, problem, or hazardous drinking, use, or consumption. We classified these alcohol consumption patterns as risky drinking if they corresponded to the definition given in **Table 1**. To be included in the systematic review, all reference standards had to have been obtained by using formal instruments that were either clearly described or referenced.

Data Extraction and Quality Assessment

We used only published data. Two reviewers extracted data on study characteristics and outcomes according to a structured form. A third reviewer double-checked extracted data. When results were reported for more than 1 cutoff point, we followed the recommendation of the primary investigators. In the case of multiple or missing recommendations, we chose the cutoff point with the highest value for the term sensitivity + specificity − 1 (the Youden index) (24). To assess clinical heterogeneity, we summarized study characteristics by using descriptive statistical methods.

One reviewer assessed methodological heterogeneity by using a methodological quality checklist derived from the QUADAS (Quality Assessment of Studies of Diagnos-

Table 1. **Definition of Target Conditions and Reference Standards**

Target Condition	Definition
Risky drinking*	Consumption of alcohol amounts above a recommended level or heavy episodic (binge) drinking in the past 12 months, which places the drinker at risk for consequential damage
Harmful drinking	Clinically significant impairments or consequences of alcohol consumption in the past 12 months, as defined in the International Classification of Diseases
Alcohol abuse	Clinically significant impairments or consequences of alcohol consumption in the past 12 months, as defined in the *Diagnostic and Statistical Manual of Mental Disorders*
Alcohol dependence*	Fulfillment of ≥3 criteria that describe a chronic, maladaptive pattern of use and impairment or distress associated with alcohol consumption in the past 12 months, as defined in the International Classification of Diseases and the *Diagnostic and Statistical Manual of Mental Disorders*
Alcohol use disorder*	Fulfilling criteria for harmful drinking, alcohol abuse, or alcohol dependence
Unhealthy alcohol use*	Fulfilling criteria for risky drinking or any alcohol use disorder

* Used as a reference standard in the review.

tic Accuracy included in Systematic Reviews) tool (25). Of the 11 items in this tool, 4 had to have been fulfilled for inclusion in the review. The other items assessed whether the following criteria were met: representative spectrum of patients, acceptable delay between tests, reference standard results blinded, index test results blinded, relevant clinical information available when the test is used in practice, uninterpretable results reported, and withdrawals explained.

Data Synthesis and Analysis

We performed data synthesis according to guidelines on reviews of diagnostic accuracy (26–30). We analyzed diagnostic accuracy independently for each of the 4 reference standards. In addition, we stratified summaries by study setting to reduce clinical heterogeneity.

The recommendations regarding the best threshold score for the AUDIT and the AUDIT-C vary considerably among primary studies according to target condition, country, setting, patient characteristics, and tradeoff between sensitivity and specificity (16–18). We needed to use adjusted statistical methods because the resulting threshold effects may have made straightforward pooling of sensitivities and specificities misleading (30). We used hierarchical summary receiver-operating characteristic (ROC) curve analysis, which models accuracy, threshold effect, and dependence of accuracy on threshold (31–33), and applied a random-effects model to account for both within-study and between-study variability.

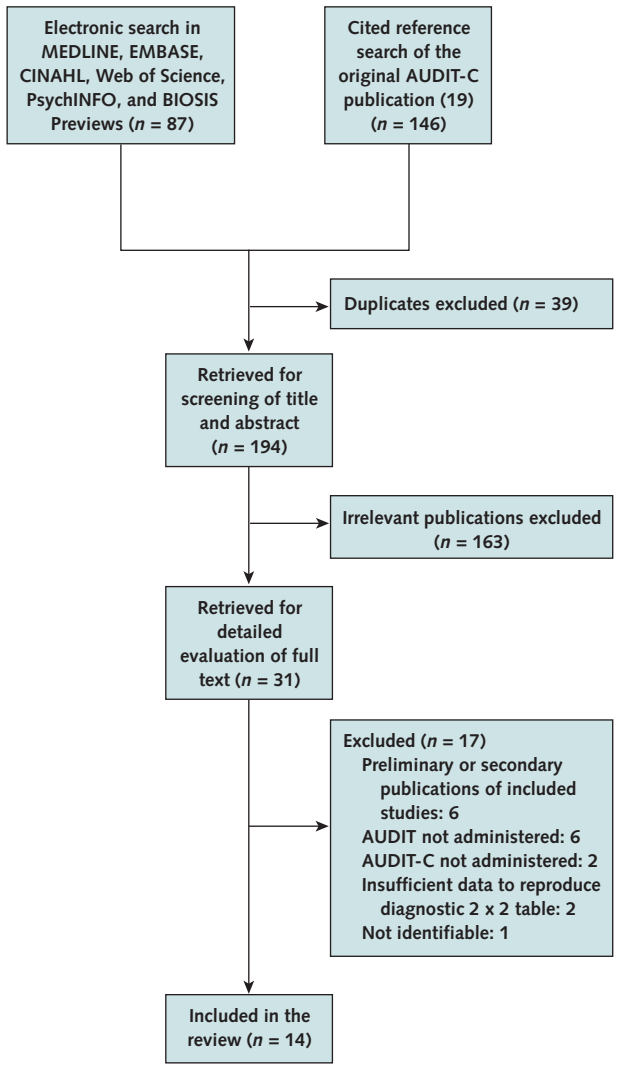
We performed further analyses to examine sex- and age-specific variability as possible effect-moderating factors in primary study results on the basis of empirical findings (17). In addition, to test for possible sample size effects that may indicate publication bias, we used effective sample size funnel plots and associated weighted regression tests for asymmetry (34). We performed statistical pooling (meta-analysis) and testing for small-study effects only when at least 4 primary studies were available. In addition, we summarized data qualitatively.

We performed hierarchical summary ROC curve analyses with Stata, version 10 (StataCorp, College Station, Texas), using the metandi command (32, 35) and created the figures with Review Manager, version 5.0 (The Cochrane Collaboration, Copenhagen, Denmark), and SigmaPlot, version 8.02 (Systat Software, San Jose, California).

Role of the Funding Source

This study was supported in part by a research grant (01EB0111) from the German Federal Ministry of Education and Research. The funding source had no influence on decisions regarding design and conduct of the study; collection, management, analysis, and interpretation of the data; or preparation and submission of the manuscript for publication.

Figure 1. **Study flow diagram.**



AUDIT = Alcohol Use Disorders Identification Test; AUDIT-C = Alcohol Use Disorders Identification Test–Consumption.

RESULTS

Retrieval of Studies

Electronic database searches and the cited reference search of the original AUDIT-C publication (19) yielded 194 references (**Figure 1**). After screening the titles and abstracts, we retrieved the full texts of 31 publications for detailed evaluation. We included 14 studies in the systematic review (19–21, 36–46). Reviewers disagreed about 5 (16.1%) of 31 studies, all of which we ultimately included. We excluded 1 study (47) because it seemed to have been performed in a subsample of another included study (40).

Characteristics of Included Studies

Appendix Table 1 (available at www.annals.org) describes the included studies. More than half (8 of 14) of the investigations were published in the last 3 years. Eight

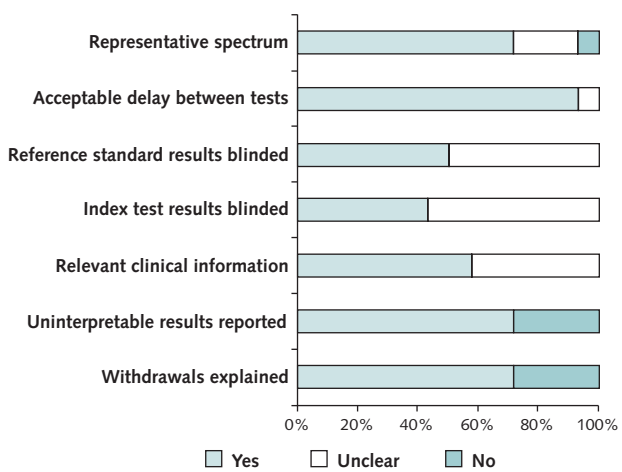
studies were done in Europe, 5 in the United States, and 1 in China. Most research (8 studies) was performed in primary care. General population samples were examined in 4 studies and inpatient samples in 2 studies. We identified no study that was performed in an emergency department. Reported sample size ranged from 112 to 13 438 persons (median, 609 persons). Sex and age distributions varied considerably.

Six of the 11 studies investigating screening accuracy for the detection of risky drinking defined this target condition by quantity or frequency of consumption and heavy episodic drinking. The standardized threshold for risky drinking varied from 196 to 280 g of pure alcohol per week for men and from 98 to 169 g for women. The definition of heavy episodic drinking ranged from the consumption of 4 to 6 drinks on a single occasion for men and from 3 to 4 drinks for women. All studies used inquiries in terms of quantity and frequency of alcohol consumption to establish the reference standard of risky drinking. In all but 1 case, trained interviewers performed interview procedures. Seven of the 9 studies used standardized instruments.

Seven of the 9 studies examining the detection of alcohol use disorders (harmful drinking, alcohol abuse, or alcohol dependence) used the criteria of the *Diagnostic and Statistical Manual of Mental Disorders* (Third Edition Revised or Fourth Edition) to diagnose the target condition. Investigators in 2 studies applied criteria from the International Classification of Diseases, Tenth Edition. Eight of the 9 studies applied standardized instruments, administered by trained interviewers, to establish the reference diagnosis.

Figure 2 summarizes the methodological quality of the included studies; **Appendix Table 2** (available at www.annals.org) provides a detailed description. In 10 of the 14 studies, consecutive or random patient selection ensured that the sample spectrum was representative of the examined population. In 13 cases, the delay between reference and index tests proved to be very short, thus excluding a high probability of patients being treated between tests. Only about half of the studies reported blinding of reference or index test results, similar to the number that described the availability of relevant clinical information when interpreting index test results. Ten publications provided information on uninterpretable test results and withdrawals. Half of the included studies fulfilled at least 6 of the 7 criteria for assessing methodological quality.

Figure 2. **Reviewer judgments about methodological quality items, reported as percentages across all included studies.**



Appendix Table 3 (available at www.annals.org) presents further primary study information on recommended threshold scores and findings. **Appendix Figure 1** (available at www.annals.org) shows ROC curves from all studies reporting results for at least 2 threshold scores.

Findings in Primary Care

Five studies comprising 8679 patients (19, 40–42, 44) examined the screening accuracy of the AUDIT and the AUDIT-C for detecting risky drinking in primary care. The prevalence of risky drinking ranged from 10.8% to 35.4%. Threshold scores varied between 4 and 8 for the AUDIT and between 3 and 5 for the AUDIT-C. Accuracy proved to depend significantly on choice of threshold, with effects showing in opposing directions for the 2 tests (**Table 2**); thus, summary ROC curves of the 2 instruments did not run parallel to one another (**Figure 3**, *A*). When sensitivity increased to greater than 0.85, specificity decreased to a lower extent for the AUDIT-C than for the AUDIT. However, when specificity increased to greater than 0.83, sensitivity decreased to a lower extent for the AUDIT than for the AUDIT-C. Summary point estimates pointed in the same direction, with a higher sensitivity for the AUDIT-C and a higher specificity for the AUDIT (**Table 3** and **Appendix Figure 2**, *A* [available at www.annals.org]). The pooled positive likelihood ratio for the AUDIT was more than twice as high as that for the AUDIT-C, with nonoverlapping 95% CIs (**Table 3**). Considerable heterogeneity was present between study findings on each test, especially for the AUDIT-C. Four of the 5 included studies fulfilled fewer than 3 of the 7 methodological quality criteria (**Appendix Table 2**).

Five studies comprising 4572 patients (19, 20, 38, 39, 44) examined the accuracy of the AUDIT and the AUDIT-C for detecting any alcohol use disorder (harmful drinking, alcohol abuse, or alcohol dependence). The prevalence of any alcohol use disorder ranged from 8.9% to 21.8%. Threshold scores were between 2 and 8 for the AUDIT and between 2 and 5 for the AUDIT-C. The wide and overlapping prediction regions of the summary ROC curves indicate considerable heterogeneity among study findings (**Figure 3**, *B*). Point estimates of sensitivity, specificity, and likelihood ratios were generally identical (**Table 3** and **Appendix Figure 2**, *B*). Four of the 5 studies were

Table 2. **Regression Parameter Estimates of Hierarchical Summary Receiver-Operating Characteristic Curve Analysis**

Target Condition	Accuracy (95% CI)	Threshold (95% CI)	Dependence of Accuracy on Threshold (95% CI)	Relative Accuracy of Tests (95% CI)*	Significance Level of Accuracy Difference between Tests (*P* Value)
Risky drinking in primary care†					
AUDIT	3.46 (3.25 to 3.67)	0.29 (−0.34 to 0.92)	0.63 (0.05 to 1.21)	1.10 (0.58 to 2.08)	0.78
AUDIT-C	3.37 (2.76 to 3.97)	0.54 (−0.33 to 1.40)	−0.89 (−1.50 to −0.29)		
Alcohol use disorder in primary care†					
AUDIT	3.09 (2.51 to 3.67)	−0.39 (−1.14 to 0.37)	−0.74 (−1.57 to 0.08)	1.29 (0.61 to 2.73)	0.50
AUDIT-C	2.83 (2.36 to 3.31)	0.12 (−0.65 to 0.89)	−0.01 (−0.66 to 0.64)		
Unhealthy alcohol use in primary care‡					
AUDIT	3.06 (2.70 to 3.71)	−0.18 (−1.30 to 0.95)	−0.13 (−1.54 to 1.29)	0.73 (0.44 to 1.22)	0.23
AUDIT-C	3.37 (3.00 to 3.74)	0.73 (0.03 to 1.41)	0.56 (0.06 to 1.07)		

AUDIT = Alcohol Use Disorders Identification Test; AUDIT-C = Alcohol Use Disorders Identification Test–Consumption.
* Values >1 indicate superiority of the AUDIT; values <1 indicate superiority of the AUDIT-C.
† Based on 5 studies.
‡ Based on 4 studies.

conducted in the United States, and all but 1 fulfilled at least 6 of the 7 methodological quality criteria (**Appendix Table 2**).

No study performed in primary care investigated the accuracy of the AUDIT and the AUDIT-C for detecting alcohol dependence alone.

Four studies comprising 2580 patients (19, 20, 39, 44) investigated the screening accuracy of the AUDIT and the AUDIT-C for detecting unhealthy alcohol use (risky drinking or any alcohol use disorder). Prevalence ranged from 22.6% to 41.2%, and threshold scores varied between 3 and 5 for the AUDIT and between 2 and 4 for the AUDIT-C. Summary ROC curves ran generally parallel with overlapping prediction regions (**Figure 3**, *C*), although moderate heterogeneity among study findings was present. Regression parameter estimates suggest that the AUDIT is slightly but nonsignificantly less accurate than the AUDIT-C, with accuracy significantly dependent on choice of threshold for the AUDIT-C (**Table 2**). Summary point estimates were fairly similar, with overlapping CIs, although the AUDIT shows a slightly lower sensitivity, a higher specificity, and a higher positive likelihood ratio than the AUDIT-C (**Table 3** and **Appendix Figure 2**, *C*). All included studies were performed in the United States, and 3 of 4 fulfilled at least 6 of 7 methodological quality criteria (**Appendix Table 2**).

On the basis of the Bayes theorem and the summary point estimates of likelihood ratios presented in **Table 3**, we calculated predictive values (posttest probabilities) on the basis of the target conditions and their prevalence rates (pretest probabilities). These values may be helpful in assessing the effects of the choice between the 2 instruments in clinical practice (**Table 4**).

Findings in Other Settings

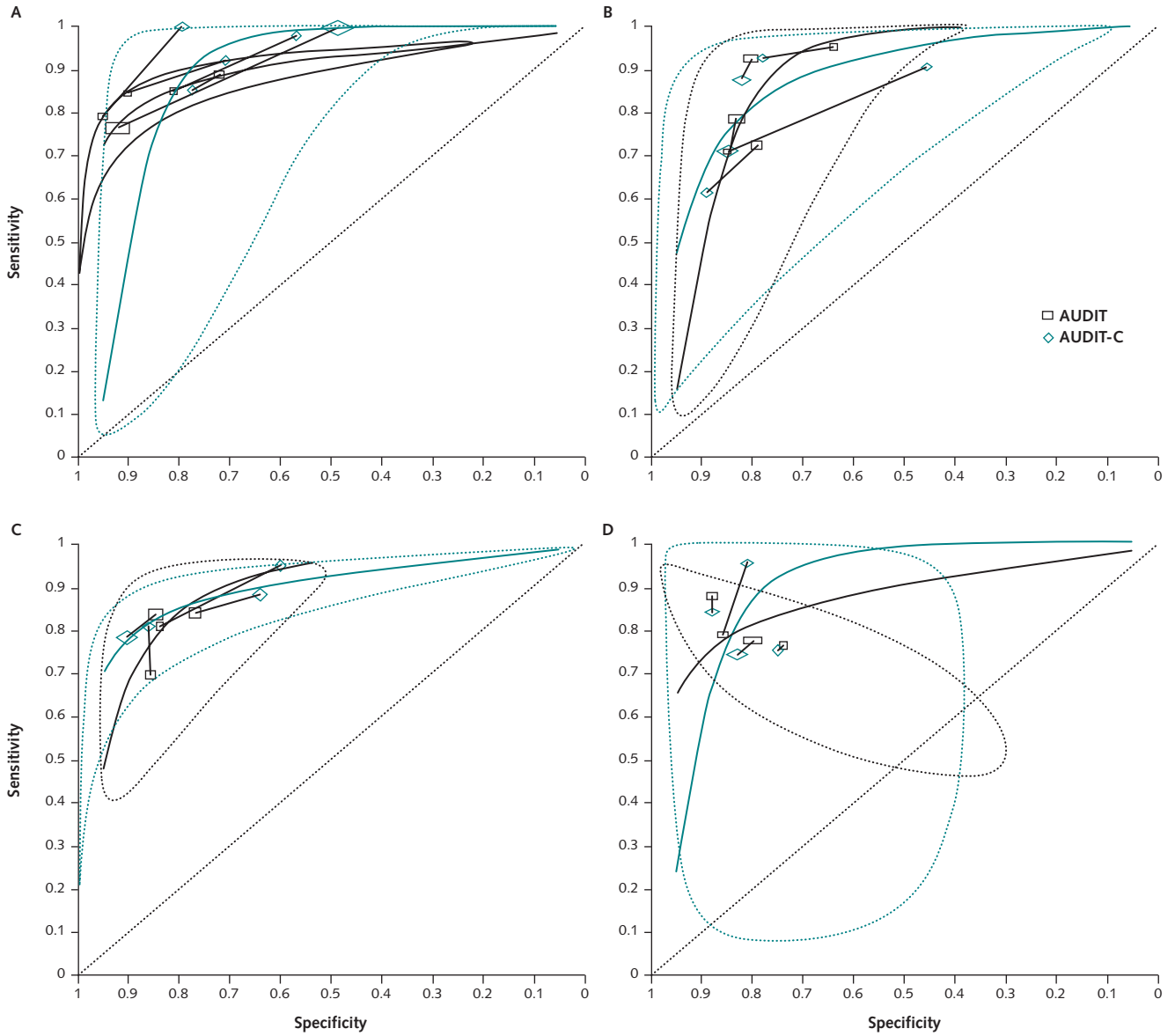
Four studies comprising 5600 adults (21, 36, 43, 46) investigated the accuracy of the AUDIT and the AUDIT-C in general population samples. The prevalence of risky drinking ranged from 5.4% to 36.5%, and threshold scores ranged from 5 to 8 for the AUDIT and from 5

Table 3. **Summary Point Estimates of Hierarchical Summary Receiver-Operating Characteristic Curve Analysis**

Target Condition	Sensitivity (95% CI)	Specificity (95% CI)	Positive Likelihood Ratio (95% CI)	Negative Likelihood Ratio (95% CI)
Risky drinking in primary care*				
AUDIT	0.81 (0.75–0.86)	0.88 (0.79–0.93)	6.62 (4.11–10.69)	0.21 (0.17–0.27)
AUDIT-C	0.97 (0.90–0.99)	0.68 (0.56–0.77)	2.99 (2.22–4.03)	0.04 (0.01–0.14)
Alcohol use disorder in primary care*				
AUDIT	0.84 (0.71–0.92)	0.79 (0.73–0.84)	4.03 (3.26–4.98)	0.20 (0.11–0.37)
AUDIT-C	0.82 (0.69–0.91)	0.78 (0.65–0.88)	3.82 (2.46–5.93)	0.22 (0.13–0.37)
Unhealthy alcohol use in primary care†				
AUDIT	0.81 (0.75–0.85)	0.83 (0.79–0.87)	4.82 (3.97–5.85)	0.23 (0.18–0.29)
AUDIT-C	0.86 (0.79–0.91)	0.78 (0.62–0.89)	3.91 (2.25–6.80)	0.18 (0.13–0.24)

AUDIT = Alcohol Use Disorders Identification Test; AUDIT-C = Alcohol Use Disorders Identification Test–Consumption.
* Based on 5 studies.
† Based on 4 studies.

Figure 3. **Summary receiver-operating characteristic curves for the Alcohol Use Disorders Identification Test (*AUDIT*) and the Alcohol Use Disorders Identification Test–Consumption (*AUDIT-C*).**



Dotted lines indicate 95% prediction regions and pairs of connected points indicate single studies. **A.** Accuracy for risky drinking in primary care. **B.** Accuracy for alcohol use disorders in primary care. **C.** Accuracy for unhealthy alcohol use in primary care. **D.** Accuracy for risky drinking in the general population.

to 6 for the AUDIT-C. Panel D of **Figure 3** shows very large heterogeneity between studies and suggests that the applied statistical model failed to fit the data. In 1 of the 4 studies (36), the AUDIT showed a slightly higher sensitivity than the AUDIT-C (0.87 vs. 0.84) with the same specificity (0.88). Two studies (43, 46) reported similar sensitivities (about 0.75) for both tests, whereas specificities were similar within studies but differed between investigations (0.80 vs. 0.74 for the AUDIT and 0.83 vs. 0.75 for the AUDIT-C). One study (21) found a higher sensitivity for the AUDIT-C than for the AUDIT (0.95 vs. 0.78) but a lower specificity (0.81 vs. 0.86). The accuracy of the instruments in detecting any alcohol use disorder was not investigated in any of the general population studies. In 1 of the 2 studies investigating accuracy for detecting alcohol dependence (21), the AUDIT showed slightly higher specificity (0.85) than the AUDIT-C (0.78) at the same sensitivity (0.79). The other study (43) yielded both higher sensitivity and specificity for the AUDIT (0.92 and 0.88) than for the AUDIT-C (0.88 and 0.81). The only study

that examined the detection of unhealthy alcohol use reported a sensitivity of 0.78 and a specificity of 0.81 for the AUDIT and a sensitivity of 0.74 and a specificity of 0.85 for the AUDIT-C (43). All research in general population samples was done in Europe, and 3 of the 4 studies fulfilled at least 5 of the 7 methodological quality criteria (**Appendix Table 2**).

Two studies, 1 in Belgium (37) and 1 in China (45), analyzed data from 345 inpatients. Neither study examined the detection of risky drinking or unhealthy alcohol use in general among these patients. Both studies investigated screening accuracy for any alcohol use disorder; 1 study (37) found that the AUDIT had a higher sensitivity (0.83 vs. 0.66) but the AUDIT-C had a higher specificity (0.94 vs. 0.85), whereas the other study (45) found that the AUDIT had a higher specificity than the AUDIT-C (0.85 vs. 0.73) with similar sensitivities (0.96 vs. 0.98). The latter study also investigated the accuracy of the instruments by detecting alcohol dependence alone. Results again showed slightly higher specificity of the AUDIT than the AUDIT-C (0.63 vs. 0.58) at the same sensitivity (0.94) (45).

We found no study that compared the accuracy of the AUDIT with that of the AUDIT-C for detecting any form of unhealthy alcohol use among patients in emergency departments.

Sex- and Age-Specific Findings

Three studies were done in men only (19, 37, 46) and 2 in women only (20, 36). In 3 further studies (38, 39, 42), analyses were stratified for sex. The prevalence of all investigated target conditions was higher for men than for women. Recommended threshold scores were also higher for men (AUDIT, 5 to 8; AUDIT-C, 3 to 6) than for women (AUDIT, 2 to 6; AUDIT-C, 2 to 5), with higher threshold scores as the severity of the target condition increased. We could not detect any systematic variation in accuracy between the AUDIT and the AUDIT-C due to sex distribution of the samples.

One study (19) was performed in a sample with a mean age above 65 years and another (40) reported findings that were stratified for age. For elderly patients, threshold scores ranged from 5 to 8 for the AUDIT and were 3 in both studies of the AUDIT-C. Age did not affect the outcome of the comparisons of the 2 instruments.

Tests for Publication Bias

Regression tests of accuracy against effective sample size did not reveal any statistically significant small-study effect (**Appendix Table 4** and **Appendix Figure 3**, available at www.annals.org), although visual examination of funnel plots indicated that smaller studies that investigate the general population may result in higher diagnostic accuracy (**Appendix Figure 3**, *D*).

Discussion

We identified 14 studies that directly compared the AUDIT with the AUDIT-C. Primary care is the most intensively researched setting, with most activities focused on screening for risky drinking, alcohol use disorders, and unhealthy alcohol use. Accuracy among patients in the general and inpatient populations has not been investigated in the United States, and we identified no studies conducted in patients in emergency departments. The methodological quality of the included studies varied considerably; the most apparent problem was nonreporting of blinding of reference or index test results, although this is recommended in the relevant guidelines (48). Recommended threshold scores for the instruments depended on sex and varied noticeably, probably because of a subjective tradeoff between sensitivity and specificity.

In quantitative analyses, the accuracy of the AUDIT and the AUDIT-C did not significantly differ for screening for risky drinking, alcohol use disorders, or unhealthy alcohol use in primary care. However, this finding is based on small sets of heterogeneous studies and should not be

Table 4. **Predictive Value Estimates Based on Likelihood Ratio Estimates of Hierarchical Summary Receiver-Operating Characteristic Curve Analysis**

Target Condition	Positive Predictive Value, by Prevalence								Negative Predictive Value, by Prevalence							
	5%	10%	15%	20%	25%	30%	40%	50%	5%	10%	15%	20%	25%	30%	40%	50%
Risky drinking in primary care																
AUDIT	0.26	0.42	0.54	0.62	0.69	0.74	0.82	0.87	0.99	0.98	0.96	0.95	0.93	0.92	0.88	0.83
AUDIT-C	0.14	0.25	0.35	0.43	0.50	0.56	0.67	0.75	1.00	1.00	0.99	0.99	0.99	0.98	0.97	0.96
Alcohol use disorder in primary care																
AUDIT	0.17	0.31	0.42	0.50	0.57	0.63	0.73	0.80	0.99	0.98	0.97	0.95	0.94	0.92	0.88	0.83
AUDIT-C	0.17	0.30	0.40	0.49	0.56	0.62	0.72	0.79	0.99	0.98	0.96	0.95	0.93	0.91	0.87	0.82
Unhealthy alcohol use in primary care																
AUDIT	0.20	0.35	0.46	0.55	0.62	0.67	0.76	0.83	0.99	0.98	0.96	0.95	0.93	0.91	0.87	0.81
AUDIT-C	0.17	0.30	0.41	0.49	0.57	0.63	0.72	0.80	0.99	0.98	0.97	0.96	0.94	0.93	0.89	0.85

AUDIT = Alcohol Use Disorders Identification Test; AUDIT-C = Alcohol Use Disorders Identification Test–Consumption.

considered as evidence for equivalence of the 2 tests (49). Some results, such as the large difference between positive likelihood ratios when screening for risky drinking in primary care, suggest that the AUDIT may be superior to the AUDIT-C. Correspondingly, in qualitative analyses, the AUDIT proved to be slightly more accurate than the AUDIT-C for screening for any alcohol use disorder or alcohol dependence alone in samples of the general population or inpatients. Because no widely accepted consensus has been reached on the definition of a "clinically relevant difference" between the accuracy of 2 screening tests (and it may depend on the setting in which it is applied), interpretation of our findings may vary. In addition to possible accuracy differences, we found no evidence that the choice between the 2 instruments should depend on either the sex or age of the target population. We also did not identify any statistically significant small-study effect that would suggest publication bias.

Although we stratified analyses for target condition and setting and examined possible sex- and age-specific variability as possible effect-moderating factors, statistical heterogeneity remained considerable for most analyses. One reason for the consistently larger heterogeneity among results for the AUDIT-C may be its lower reliability compared with that of the full AUDIT (17). Another source of heterogeneity may be the variation in the reference standards, especially for risky drinking. However, this is inevitable because of population-specific thresholds for low-risk alcohol consumption (for example, the recommended maximum amount is usually lower for women than for men). Sources of this variability should be investigated in further research. Some of the observed heterogeneity may be explained by such factors as ethnicity, cultural background, education level of the patients in the sample, the extent of patients' acceptance of having a drinking problem, the professional background and degree of training of those administering the index and reference tests, and methodological issues.

Apart from the AUDIT, the Cut-Down, Annoyed, Guilty, and Eye-opener (CAGE) questionnaire (50) and the Michigan Alcoholism Screening Test (MAST) (51) are the most frequently recommended and researched case-finding instruments for unhealthy alcohol use. Further comparative reviews would be needed to resolve the equivalency of these tests. However, several features beyond statistical screening accuracy may influence the choice of an instrument. Even if the AUDIT-C is shorter than the full AUDIT, the latter may serve as a starting point for the exploration of the alcohol problem in a general practice situation because of its questions about the consequences of alcohol use. The MAST can provide a detailed description of a potential alcohol problem in settings where time constraints are not crucial. Finally, the CAGE instrument, with its 4 easily memorizable yes-or-no questions, may be preferable to both the AUDIT and the AUDIT-C, which have several response categories.

An optimal metric to summarize findings on instruments with several possible thresholds may focus on the misclassification rates at the high- and low-risk ends of the diagnostic continuum by comparing predictiveness curves for the 2 instruments. A predictiveness curve plots the risk for a target condition against the test value quantiles in a certain population (52, 53). This procedure transforms the original scales to a common scale (quantiles), thus enabling comparisons between tests that are not possible by using their original scales. In addition, this method does not depend on the choice of threshold by the primary investigators (which may show uncontrollable variation) but rather relies on the reviewers' definitions of "high risk" and "low risk" (which may be held constant for comparisons). For paired studies, differences in proportions of persons correctly classified as high or low risk between 2 tests may be calculated and pooled by using standard meta-analytic techniques. The curves themselves may also be pooled, which may be of particular relevance for comparisons between tests that were examined in different sets of studies. This procedure would require that investigators report the distribution of the target condition at all possible thresholds. Thus, the question of noninferiority of the AUDIT-C to the full AUDIT is most likely to be resolved through individual patient data–based collaborative meta-analysis.

Until such evidence is available and further advances are made in the interpretation of meta-analyzed screening and diagnostic test characteristics, the issue of whether 3 questions are enough to detect unhealthy alcohol use remains without a definite answer.

From the University Medical Center Freiburg, Freiburg, and University Medical Center Hamburg-Eppendorf, Hamburg, Germany.

Acknowledgment: The authors thank Sarah Mannion for copyediting the manuscript.

Grant Support: In part by a research grant (01EB0111) from the German Federal Ministry of Education and Research.

Potential Financial Conflicts of Interest: None disclosed.

Requests for Single Reprints: Levente Kriston, MA, Section Clinical Epidemiology and Health Services Research, Department of Psychiatry and Psychotherapy, University Medical Center Freiburg, Hauptstrasse 5, D-79104, Freiburg, Germany; e-mail, levente.kriston@uniklinik-freiburg.de.

Current author addresses are available at www.annals.org.

References

1. **Hasin DS, Stinson FS, Ogburn E, Grant BF.** Prevalence, correlates, disability, and comorbidity of DSM-IV alcohol abuse and dependence in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. Arch Gen Psychiatry. 2007;64:830-42. [PMID: 17606817]

2. **Reid MC, Fiellin DA, O'Connor PG.** Hazardous and harmful alcohol consumption in primary care. Arch Intern Med. 1999;159:1681-9. [PMID: 10448769]

3. **Watson H.** Problem drinkers among acute care inpatients. Nurs Stand. 2000; 14:32-5. [PMID: 11974176]
4. **Orford J, Somers M, Daniels V, Kirby B.** Drinking amongst medical patients: levels of risk and models of change. Br J Addict. 1992;87:1691-702. [PMID: 1490083]
5. **Dunn CW, Donovan DM, Gentilello LM.** Practical guidelines for performing alcohol interventions in trauma centers. J Trauma. 1997;42:299-304. [PMID: 9042886]
6. **D'Onofrio G, Bernstein E, Bernstein J, Woolard RH, Brewer PA, Craig SA, et al.** Patients with alcohol problems in the emergency department, part 1: improving detection. SAEM Substance Abuse Task Force. Society for Academic Emergency Medicine. Acad Emerg Med. 1998;5:1200-9. [PMID: 9864134]
7. **Saitz R.** Clinical practice. Unhealthy alcohol use. N Engl J Med. 2005;352: 596-607. [PMID: 15703424]
8. **Harwood HJ, Fountain D, Fountain G.** Economic cost of alcohol and drug abuse in the United States, 1992: a report. Addiction. 1999;94:631-5. [PMID: 10563025]
9. **Rice DP.** Economic costs of substance abuse, 1995. Proc Assoc Am Physicians. 1999;111:119-25. [PMID: 10220806]
10. **Gmel G, Rehm J.** Harmful alcohol use. Alcohol Res Health. 2003;27:52-62. [PMID: 15301400]
11. **Babor TF, Higgins-Biddle JC.** Brief Intervention for Hazardous and Harmful Drinking. A Manual for Use in Primary Care. Geneva: World Health Organization; 2001.
12. Helping Patients Who Drink Too Much. A Clinician's Guide. Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism; 2005.
13. The Management of Harmful Drinking and Alcohol Dependence in Primary Care. A National Clinical Guideline. Edinburgh, Scotland: Scottish Intercollegiate Guidelines Network; 2003.
14. **Fiellin DA, Reid MC, O'Connor PG.** Screening for alcohol problems in primary care: a systematic review. Arch Intern Med. 2000;160:1977-89. [PMID: 10888972]
15. **Maisto SA, Saitz R.** Alcohol use disorders: screening and diagnosis. Am J Addict. 2003;12 Suppl 1:S12-25. [PMID: 14972777]
16. **Babor TF, Higgins-Biddle JC, Saunders JB, Monteiro MG. AUDIT.** The Alcohol Use Disorders Identification Test. Guidelines for Use in Primary Health Care. 2nd ed. Geneva: World Health Organization; 2001.
17. **Reinert DF, Allen JP.** The alcohol use disorders identification test: an update of research findings. Alcohol Clin Exp Res. 2007;31:185-99. [PMID: 17250609]
18. **Berner MM, Kriston L, Bentele M, Härter M.** The alcohol use disorders identification test for detecting at-risk drinking: a systematic review and meta-analysis. J Stud Alcohol Drugs. 2007;68:461-73. [PMID: 17446987]
19. **Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA.** The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. Arch Intern Med. 1998;158:1789-95. [PMID: 9738608]
20. **Bradley KA, Bush KR, Epler AJ, Dobie DJ, Davis TM, Sporleder JL, et al.** Two brief alcohol-screening tests. From the Alcohol Use Disorders Identification Test (AUDIT): validation in a female Veterans Affairs patient population. Arch Intern Med. 2003;163:821-9. [PMID: 12695273]
21. **Selin KH.** Alcohol Use Disorder Identification Test (AUDIT): what does it screen? Performance of the AUDIT against four different criteria in a Swedish population sample. Subst Use Misuse. 2006;41:1881-99. [PMID: 17162595]
22. **Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF.** Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. QUOROM Group. Br J Surg. 2000;87:1448-54. [PMID: 11091231]
23. **Reinert DF, Allen JP.** The Alcohol Use Disorders Identification Test (AUDIT): a review of recent research. Alcohol Clin Exp Res. 2002;26:272-9. [PMID: 11964568]
24. **Youden WJ.** Index for rating diagnostic tests. Cancer. 1950;3:32-5. [PMID: 15405679]
25. **Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J.** The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol. 2003;3:25. [PMID: 14606960]
26. **Battaglia M, Bucher H, Egger M, Grossenbacher F, Minder C, Pewsner D.** The Bayes Library of Diagnostic Studies and Reviews. Bern, Switzerland: University of Bern; 2002.
27. **Deeks JJ.** Systematic reviews of evaluations of diagnostic and screening tests. In: Egger M, Smith GD, Altman DG, eds. Systematic Reviews in Health Care: Meta-Analysis in Context. 2nd ed. London: BMJ; 2003:248-82.
28. **Devillé WL, Buntinx F, Bouter LM, Montori VM, de Vet HC, van der Windt DA, et al.** Conducting systematic reviews of diagnostic studies: didactic guidelines. BMC Med Res Methodol. 2002;2:9. [PMID: 12097142]
29. **Irwig L, Tosteson AN, Gatsonis C, Lau J, Colditz G, Chalmers TC, et al.** Guidelines for meta-analyses evaluating diagnostic tests. Ann Intern Med. 1994; 120:667-76. [PMID: 8135452]
30. **Gatsonis C, Paliwal P.** Meta-analysis of diagnostic and screening test accuracy evaluations: methodologic primer. AJR Am J Roentgenol. 2006;187:271-81. [PMID: 16861527]
31. **Rutter CM, Gatsonis CA.** A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. Stat Med. 2001;20:2865-84. [PMID: 11568945]
32. **Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne JA.** A unification of models for meta-analysis of diagnostic accuracy studies. Biostatistics. 2007;8:239-51. [PMID: 16698768]
33. **Macaskill P.** Empirical Bayes estimates generated in a hierarchical summary ROC analysis agreed closely with those of a full Bayesian analysis. J Clin Epidemiol. 2004;57:925-32. [PMID: 15504635]
34. **Deeks JJ, Macaskill P, Irwig L.** The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. J Clin Epidemiol. 2005;58:882-93. [PMID: 16085191]
35. **Harbord RM.** Metandi: Stata module for meta-analysis of diagnostic accuracy. Boston: Boston Coll Dept Econ; 2008.
36. **Aalto M, Tuunanen M, Sillanaukee P, Seppä K.** Effectiveness of structured questionnaires for screening heavy drinking in middle-aged women. Alcohol Clin Exp Res. 2006;30:1884-8. [PMID: 17067353]
37. **Aertgeerts B, Buntinx F, Ansoms S, Fevery J.** Questionnaires are better than laboratory tests to screen for current alcohol abuse or dependence in a male inpatient population. Acta Clin Belg. 2002;57:241-9. [PMID: 12534130]
38. **Aertgeerts B, Buntinx F, Ansoms S, Fevery J.** Screening properties of questionnaires and laboratory tests for the detection of alcohol abuse or dependence in a general practice population. Br J Gen Pract. 2001;51:206-17. [PMID: 11255902]
39. **Bradley KA, DeBenedetti AF, Volk RJ, Williams EC, Frank D, Kivlahan DR.** AUDIT-C as a brief screen for alcohol misuse in primary care. Alcohol Clin Exp Res. 2007;31:1208-17. [PMID: 17451397]
40. **Gómez A, Conde A, Santana JM, Jorrín A, Serrano IM, Medina R.** The diagnostic usefulness of AUDIT and AUDIT-C for detecting hazardous drinkers in the elderly. Aging Ment Health. 2006;10:558-61. [PMID: 16938691]
41. **Gordon AJ, Maisto SA, McNeil M, Kraemer KL, Conigliaro RL, Kelley ME, et al.** Three questions can detect hazardous drinkers. J Fam Pract. 2001;50: 313-20. [PMID: 11300982]
42. **Gual A, Segura L, Contel M, Heather N, Colom J.** Audit-3 and audit-4: effectiveness of two short forms of the alcohol use disorders identification test. Alcohol Alcohol. 2002;37:591-6. [PMID: 12414553]
43. **Rumpf HJ, Hapke U, Meyer C, John U.** Screening for alcohol use disorders and at-risk drinking in the general population: psychometric performance of three questionnaires. Alcohol Alcohol. 2002;37:261-8. [PMID: 12003915]
44. **Seale JP, Boltri JM, Shellenberger S, Velasquez MM, Cornelius M, Guyinn M, et al.** Primary care validation of a single screening question for drinkers. J Stud Alcohol. 2006;67:778-84. [PMID: 16847548]
45. **Tsai MC, Tsai YF, Chen CY, Liu CY.** Alcohol Use Disorders Identification Test (AUDIT): establishment of cut-off scores in a hospitalized Chinese population. Alcohol Clin Exp Res. 2005;29:53-7. [PMID: 15654291]
46. **Tuunanen M, Aalto M, Seppä K.** Binge drinking and its detection among middle-aged men using AUDIT, AUDIT-C and AUDIT-3. Drug Alcohol Rev. 2007;26:295-9. [PMID: 17454019]
47. **Gómez A, Conde A, Santana JM, Jorrín A.** Diagnostic usefulness of brief versions of Alcohol Use Disorders Identification Test (AUDIT) for detecting hazardous drinkers in primary care settings. J Stud Alcohol. 2005;66:305-8. [PMID: 15957683]
48. **Standards for Reporting of Diagnostic Accuracy.** Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD Initiative. Ann Intern Med. 2003;138:40-4. [PMID: 12513043]
49. **Altman DG, Bland JM.** Absence of evidence is not evidence of absence. BMJ. 1995;311:485. [PMID: 7647644]

50. **Ewing JA.** Detecting alcoholism. The CAGE questionnaire. JAMA. 1984; 252:1905-7. [PMID: 6471323]
51. **Selzer ML.** The Michigan alcoholism screening test: the quest for a new diagnostic instrument. Am J Psychiatry. 1971;127:1653-8. [PMID: 5565851]
52. **Huang Y, Sullivan Pepe M, Feng Z.** Evaluating the predictiveness of a continuous marker. Biometrics. 2007;63:1181-8. [PMID: 17489968]
53. **Pepe MS, Feng Z, Huang Y, Longton G, Prentice R, Thompson IM, et al.** Integrating the predictiveness of a marker with its performance as a classifier. Am J Epidemiol. 2008;167:362-8. [PMID: 17982157]

Current Author Addresses: Mr. Kriston, Mr. Hölzel, Ms. Weiser, and Dr. Berner: Section Clinical Epidemiology and Health Services Research, Department of Psychiatry and Psychotherapy, University Medical Center Freiburg, Hauptstrasse 5, D-79104, Freiburg, Germany.
Dr. Härter: Department of Medical Psychology, University Medical Center Hamburg-Eppendorf, Martinistrasse 52, D-20246, Hamburg, Germany.

Appendix Table 1. **Study Characteristics**

Study, Year (Reference)	Country	Setting	Patients (Analyzed), *n* (*n*)	Men, %	Age, *y*	Target Conditions	Definition	Instrument; Method	Funding Source
Aalto et al., 2006 (36)	Finland	General population	971 (894)	0	40	Risky drinking	≥140 g alcohol/wk	Timeline followback; in-person interview by nurses	NR
Aertgeerts et al., 2001 (38)	Belgium	Primary care	2073 (1992)	48.8	Men, 54 (mean); women, 48 (mean)	Alcohol abuse, alcohol dependence	Active disorder according to the DSM-III-R	Composite International Diagnostic Interview; self-administered questionnaire	Merck-Belgolabo (partially); Orange Medical
Aertgeerts et al., 2002 (37)	Belgium	Inpatients	233 (233)	100	62 (mean)	Alcohol abuse, alcohol dependence	Active disorder according to the DSM-III-R	Composite International Diagnostic Interview; self-administered questionnaire	Merck-Belgolabo (partially); Orange Medical
Bradley et al., 2003 (20)	United States	Primary care	393 (393)	0	72.0% in range 30 to 59	Risky drinking	>7 drinks/wk or ≥4 drinks/occasion	Alcohol Use Disorders and Associated Disabilities Interview Schedule; in-person interview by trained nonclinician interviewers	U.S. Department of Veteran Affairs
						Alcohol abuse, alcohol dependence	Active disorder according to the DSM-IV	Alcohol Use Disorders and Associated Disabilities Interview Schedule; in-person interview by trained nonclinician interviewers	
Bradley et al., 2007 (39)	United States	Primary care	1319 (1319)	29.7	Men, 63.5% in range 30 to 64; women, 64.4% in range 30 to 64	Risky drinking	>14 drinks/wk or ≥5 drinks/occasion for men; >7 drinks/wk or ≥4 drinks/occasion for women	Quantity and frequency of consumption questions; in-person interview by trained nonclinician interviewers	National Institute on Alcohol Abuse and Alcoholism; Bureau of Health Professions, Health Resources and Services Administration
						Alcohol abuse, alcohol dependence	Active disorder according to the DSM-IV	Alcohol Use Disorders and Associated Disabilities Interview Schedule; in-person interview by trained nonclinician interviewers	
Bush et al., 1998 (19)	United States	Primary care	243 (243)	100	67 (mean)	Risky drinking	>14 drinks/wk or ≥5 drinks/occasion	World Health Organization trilevel alcohol consumption interview; telephone interview by experienced interviewers	U.S. Department of Veteran Affairs; Washington Alcohol and Drug Abuse Institute; Veteran Affairs Puget Sound Health Care System
						Alcohol abuse, alcohol dependence	Lifetime disorder according to the DSM-III-R, fulfilling ≥1 criterion in past year	Alcohol Use Disorders and Associated Disabilities Interview Schedule; telephone interview by experienced interviewers	
Gómez et al., 2006 (40)	Spain	Primary care	602 (602)	45.0	Elderly, 72 (mean); adults, 38 (mean)	Risky drinking	≥280 g alcohol/wk for men; ≥168 g alcohol/wk for women	Quantity and frequency of consumption questions; in-person interview	NR
Gordon et al., 2001 (41)	United States	Primary care	13 438 (6954)	53.0	64.7% in range 41 to 75	Risky drinking	≥16 drinks/wk for men; ≥12 drinks/wk for women	Quantity and frequency of consumption questions; self-administered questionnaire	National Institute on Alcohol Abuse and Alcoholism
Gual et al., 2002 (42)	Spain	Primary care	255 (255)	49.8	Men, 44 (mean); women, 44 (mean)	Risky drinking	≥280 g alcohol/wk for men; ≥168 g alcohol/wk for women	Systematic Interview of Alcohol Consumption; in-person interview by the general practitioner	NR
Rumpf et al., 2002 (43)	Germany	General population	3551 (3551)	50.8	41 (mean)	Risky drinking	≥30 g alcohol/d for men; ≥20 g alcohol/d for women	Munich-Composite International Diagnostic Interview; interview by trained and supervised interviewers	Federal Ministry of Education and Research
						Alcohol abuse, alcohol dependence	Active disorder according to the DSM-IV	Munich-Composite International Diagnostic Interview; interview by trained and supervised interviewers	
Seale et al., 2006 (44)	United States	Primary care	625 (625)	45.6	41 (mean)	Risky drinking	>14 drinks/wk or >4 drinks/occasion for men; >7 drinks/wk or >3 drinks/occasion for women	Timeline followback; in-person interview by a research assistant	Medcen Foundation of Macon
						Alcohol abuse, alcohol dependence	Active disorder according to the DSM-IV	Diagnostic Interview Schedule; in-person interview by a research assistant	
Selin, 2006 (21)	Sweden	General population	600 (600)	NR	NR	Risky drinking	>9.3 L alcohol/y	Graduated frequency scale; interview	NR
						Alcohol abuse, alcohol dependence	Fulfilling ≥3 of 6 criteria at least twice in the past year according to the ICD-10	ICD-10 checklist as interview	
Tsai et al., 2005 (45)	China	Inpatients	112 (112)	69.6	50 (mean)	Harmful drinking, alcohol dependence	Active disorder according to the ICD-10	Diagnostic Interview Schedule; in-person interview by a psychiatrist	National Science Council
Tuunanen et al., 2007 (46)	Finland	General population	615 (555)	100	45	Risky drinking	≥280 g alcohol/wk or ≥6 drinks/occasion	Quantity and frequency of consumption questions; in-person interview by nurses	NR

DSM-III-R = *Diagnostic and Statistical Manual of Mental Disorders* (Third Edition Revised); DSM-IV = *Diagnostic and Statistical Manual of Mental Disorders* (Fourth Edition); ICD-10 = International Classification of Diseases, Tenth Edition; NR = not reported.

Appendix Table 2. **Reviewer Judgments about Methodological Quality**

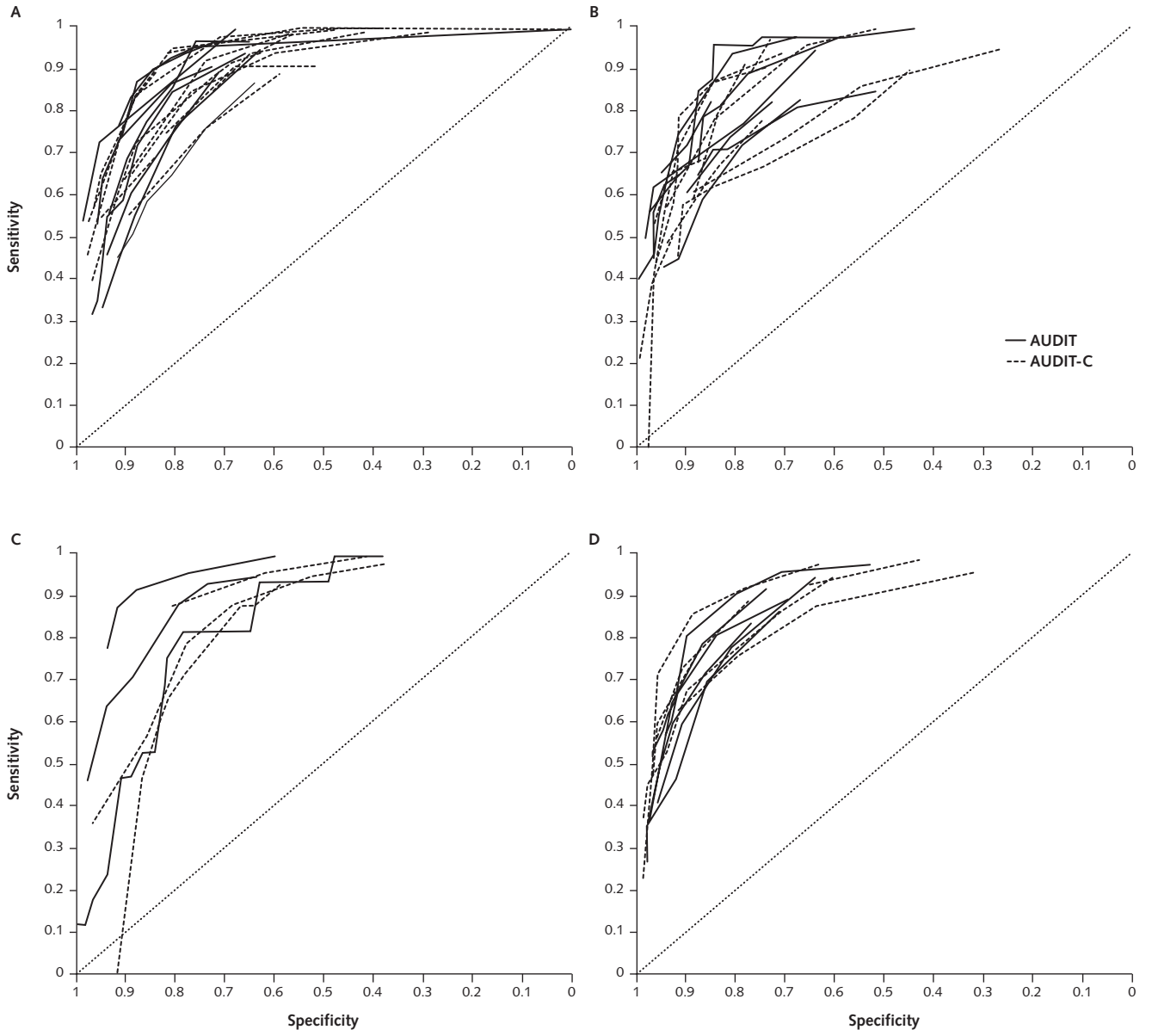
Study, Year (Reference)	Representative Spectrum of Patients	Acceptable Delay between Tests	Reference Standard Results Blinded	Index Test Results Blinded	Relevant Clinical Information Available When the Test Is Used in Practice	Uninterpretable Results Reported	Withdrawals Explained
Aalto et al., 2006 (36)	Yes	Unclear	Unclear	Unclear	Unclear	Yes	No
Aertgeerts et al., 2001 (38)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Aertgeerts et al., 2002 (37)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bradley et al., 2003 (20)	Unclear	Yes	Yes	Yes	Yes	Yes	Yes
Bradley et al., 2007 (39)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Bush et al., 1998 (19)	Unclear	Yes	Yes	Yes	Yes	Yes	Yes
Gómez et al., 2006 (40)	Yes	Yes	Unclear	Unclear	Unclear	No	No
Gordon et al., 2001 (41)	No	Yes	Unclear	Unclear	Unclear	Yes	No
Gual et al., 2002 (42)	Unclear	Yes	Unclear	Unclear	Unclear	No	Yes
Rumpf et al., 2002 (43)	Yes	Yes	Yes	Unclear	Yes	Yes	Yes
Seale et al., 2006 (44)	Yes	Yes	Unclear	Unclear	Unclear	No	No
Selin, 2006 (21)	Yes	Yes	Unclear	Unclear	Yes	Yes	Yes
Tsai et al., 2005 (45)	Yes	Yes	Yes	Unclear	Unclear	No	Yes
Tuunanen et al., 2007 (46)	Yes	Yes	Unclear	Yes	Yes	Yes	Yes

Appendix Table 3. Diagnostic Accuracy Measures in the Included Studies

Study, Year (Reference)	Subgroup	Target Condition	Prevalence, %	Test	Threshold	Sensitivity	Specificity	Positive Likelihood Ratio	Negative Likelihood Ratio	Positive Predictive Value	Negative Predictive Value
Aalto et al., 2006 (36)	All	Risky drinking	6.2	AUDIT	6*	0.87†	0.88†	7.25‡	0.14‡	0.32†	0.99†
				AUDIT-C	5*	0.84†	0.88†	6.95‡	0.19‡	0.31†	1†
Aertgeerts et al., 2001 (38)	Men	Alcohol use disorders	13.6	AUDIT	5*	0.83†	0.73†	3.05†	0.24†	0.32†	0.96†
				AUDIT-C	5*	0.78†	0.75†	3.10†	0.29†	0.33†	0.96†
	Women	Alcohol use disorders	4.5	AUDIT	5*	0.65†	0.92†	8.05†	0.35†	0.28†	0.98†
				AUDIT-C	5*	0.50†	0.93†	7.39†	0.54†	0.26†	0.98†
	All	Alcohol use disorders	8.9	AUDIT	5*	0.78‡	0.83‡	4.63‡	0.26‡	0.31‡	0.97‡
				AUDIT-C	5*	0.71‡	0.85‡	4.64‡	0.34‡	0.31‡	0.97‡
Aertgeerts et al., 2002 (37)	All	Alcohol use disorders	12.4	AUDIT	5*	0.83†	0.85†	5.63†	0.20†	0.44†	0.97†
				AUDIT-C	6§	0.66†	0.94†	10.28†	0.37†	0.59†	0.95†
Bradley et al., 2003 (20)	All	Alcohol use disorders	9.9	AUDIT	2§	0.95†	0.64†	2.60†	0.08†	0.22‡	0.99‡
				AUDIT-C	2§	0.92†	0.78†	4.20†	0.10†	0.32‡	0.99‡
	All	Unhealthy alcohol use	22.6	AUDIT	3*	0.70†	0.86†	4.90†	0.35†	0.59‡	0.91‡
				AUDIT-C	2*	0.81†	0.86†	5.90†	0.22†	0.63‡	0.94‡
Bradley et al., 2007 (39)	Men	Alcohol use disorders	16.8	AUDIT	5§	0.89†	0.78†	4.05†	0.14†	0.45‡	0.97‡
				AUDIT-C	4§	0.88†	0.75†	3.54†	0.16†	0.41‡	0.97‡
		Unhealthy alcohol use	32.7	AUDIT	4*§	0.91†	0.80†	4.55†	0.11†	0.69‡	0.95‡
				AUDIT-C	4*	0.86†	0.89†	7.82†	0.16†	0.79‡	0.93‡
	Women	Alcohol use disorders	9.0	AUDIT	3§	0.94†	0.81†	4.90†	0.07†	0.33‡	0.99‡
				AUDIT-C	3§	0.87†	0.85†	5.72†	0.16†	0.36‡	0.98‡
		Unhealthy alcohol use	19.1	AUDIT	3*	0.79†	0.87†	5.83†	0.25†	0.58‡	0.94‡
				AUDIT-C	3*	0.73†	0.91†	7.87†	0.29†	0.65‡	0.93‡
	All	Alcohol use disorders	11.3	AUDIT	SSR	0.92‡	0.80‡	4.64‡	0.10‡	0.37‡	0.99‡
				AUDIT-C	SSR	0.87‡	0.82‡	4.88‡	0.16‡	0.38‡	0.98‡
		Unhealthy alcohol use	23.1	AUDIT	SSR	0.84‡	0.85‡	5.51‡	0.19‡	0.62‡	0.95‡
				AUDIT-C	SSR	0.78‡	0.90‡	8.19‡	0.24‡	0.71‡	0.93‡
Bush et al., 1998 (19)	All	Risky drinking	35.4	AUDIT	5§	0.85†	0.81†	4.39†	0.19†	0.71‡	0.91‡
				AUDIT-C	3*	0.98†	0.57†	2.26†	0.04†	0.55‡	0.98‡
	All	Alcohol use disorders	21.4	AUDIT	8§	0.71†	0.85†	4.85†	0.34†	0.55‡	0.92‡
				AUDIT-C	3*	0.90†	0.45†	1.64†	0.21†	0.30‡	0.95‡
	All	Unhealthy alcohol use	41.2	AUDIT	5§	0.81†	0.84†	5.04†	0.23†	0.78‡	0.86‡
				AUDIT-C	3*	0.95†	0.60†	2.38†	0.08†	0.63‡	0.95‡
Gómez et al., 2006 (40)	Adults	Risky drinking	11.9	AUDIT	8*	0.83†	0.95†	16.03‡	0.17‡	0.68†	0.98‡
				AUDIT-C	3*	1†	0.79†	4.67‡	0‡	0.39†	1‡
	Elderly	Risky drinking	9.5	AUDIT	8*	0.67†	0.95†	14.25‡	0.35‡	0.60†	0.96‡
				AUDIT-C	3*	1†	0.81†	5.18‡	0‡	0.35†	1‡
	All	Risky drinking	11.1	AUDIT	8*	0.79‡	0.95‡	15.67‡	0.22‡	0.66‡	0.97‡
				AUDIT-C	3*	1‡	0.79‡	4.82‡	0‡	0.38‡	1‡
Gordon et al., 2001 (41)	All	Risky drinking	10.8	AUDIT	8*	0.76†	0.92†	9.4†	0.25†	0.54‡	0.97‡
				AUDIT-C	3*	1†	0.48†	1.9†	0.01†	0.19‡	1‡
Gual et al., 2002 (42)	Men	Risky drinking	41.7	AUDIT	7*	0.87†	0.81†	4.59‡	0.16‡	0.77†	0.90‡
				AUDIT-C	5*	0.92†	0.74†	3.60‡	0.10‡	0.72†	0.93‡
	Women	Risky drinking	8.6	AUDIT	5*	0.73†	0.96†	17.02‡	0.28‡	0.61†	0.97‡
				AUDIT-C	4*	0.91†	0.68†	2.87‡	0.13‡	0.21†	0.99‡
	All	Risky drinking	25.1	AUDIT	SSR	0.84‡	0.90‡	8.48‡	0.17‡	0.74‡	0.95‡
				AUDIT-C	SSR	0.92‡	0.71‡	3.14‡	0.11‡	0.51‡	0.96‡
Rumpf et al., 2002 (43)	All	Risky drinking	5.4	AUDIT	5*	0.77†	0.80†	3.85‡	0.29‡	0.18‡	0.98‡
				AUDIT-C	5*	0.74†	0.83†	4.34‡	0.32‡	0.20‡	0.98‡
	All	Alcohol dependence	1.4	AUDIT	6*	0.92†	0.88†	7.66‡	0.09‡	0.10‡	1‡
				AUDIT-C	5*	0.88†	0.81†	4.62‡	0.15‡	0.06‡	1‡
	All	Unhealthy alcohol use	7.9	AUDIT	5*	0.78†	0.81†	4.10‡	0.27‡	0.26‡	0.98‡
				AUDIT-C	5*	0.74†	0.85†	4.94‡	0.31‡	0.30‡	0.97‡
Seale et al., 2006 (44)	All	Risky drinking	25.7	AUDIT	4§	0.89†	0.72†	3.16‡	0.16‡	0.52‡	0.95‡
				AUDIT-C	4§	0.85†	0.77†	3.71‡	0.19‡	0.56‡	0.94‡
	All	Alcohol use disorders	21.8	AUDIT	5§	0.72†	0.79†	3.44‡	0.35‡	0.49‡	0.91‡
				AUDIT-C	5§	0.61†	0.89†	5.50‡	0.44‡	0.61‡	0.89‡
	All	Unhealthy alcohol use	35.2	AUDIT	4*	0.84†	0.77†	3.65‡	0.21‡	0.66‡	0.90‡
				AUDIT-C	3*	0.88†	0.64†	2.46‡	0.19‡	0.57‡	0.91‡
Selin, 2006 (21)	All	Risky drinking	10.0	AUDIT	8*	0.78†	0.86†	5.52†	0.25†	0.38‡	0.97‡
				AUDIT-C	6*	0.95†	0.81†	5.00†	0.06†	0.36‡	0.99‡
	All	Alcohol dependence	9.0	AUDIT	8*	0.79†	0.85†	5.15†	0.25†	0.34‡	0.98‡
				AUDIT-C	6*	0.79†	0.78†	3.56†	0.27†	0.26‡	0.97‡
Tsai et al., 2005 (45)	All	Alcohol use disorders	31.3	AUDIT	8*	0.96†	0.85†	6.41‡	0.05‡	0.85†	0.96†
				AUDIT-C	3*	0.98†	0.73†	3.68‡	0.03‡	0.76†	0.98†
	All	Alcohol dependence	15.2	AUDIT	11*	0.94†	0.63†	2.55‡	0.09‡	0.31†	0.98†
				AUDIT-C	5*	0.94†	0.58†	2.24‡	0.10‡	0.29†	0.98†
Tuunanen et al., 2007 (46)	All	Risky drinking	36.5	AUDIT	7*§	0.76†	0.74†	2.91‡	0.33‡	0.76†	0.75†
				AUDIT-C	6*	0.75†	0.75†	3.00‡	0.33‡	0.77†	0.74†

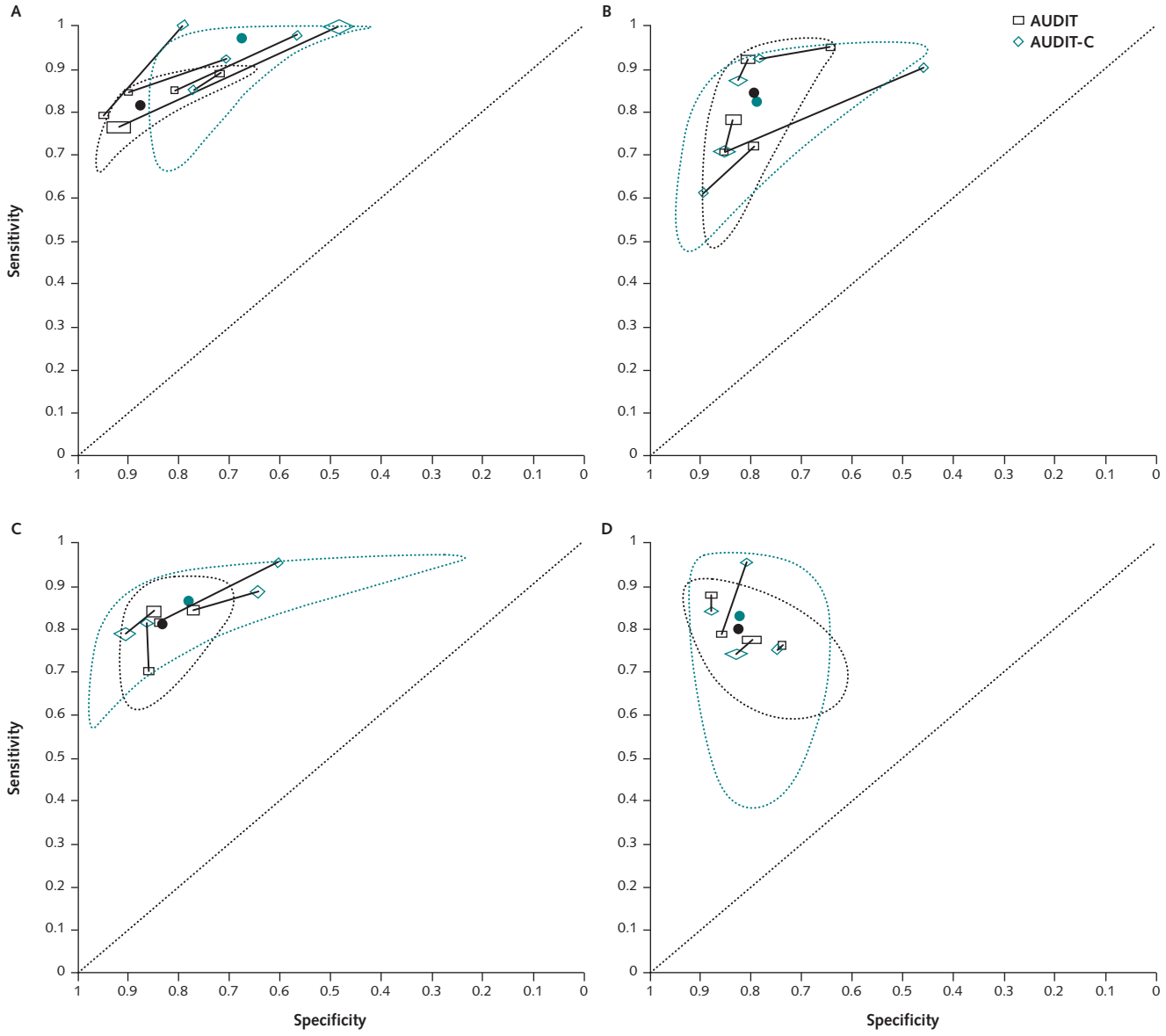
AUDIT = Alcohol Use Disorders Identification Test; AUDIT-C = Alcohol Use Disorders Identification Test–Consumption; SSR = subgroup-specific recommendation.
* Author recommendation.
† Reported in primary publication.
‡ Recalculated by the reviewers.
§ Chosen by Youden index.

Appendix Figure 1. **Receiver-operating characteristic curves from the primary studies for each reference standard.**



AUDIT = Alcohol Use Disorders Identification Test; AUDIT-C = Alcohol Use Disorders Identification Test–Consumption. **A.** Risky drinking. **B.** Alcohol use disorders. **C.** Alcohol dependence. **D.** Unhealthy alcohol use.

Appendix Figure 2. **Summary point estimates for the Alcohol Use Disorders Identification Test (*AUDIT*) and the Alcohol Use Disorders Identification Test–Consumption (*AUDIT-C*).**



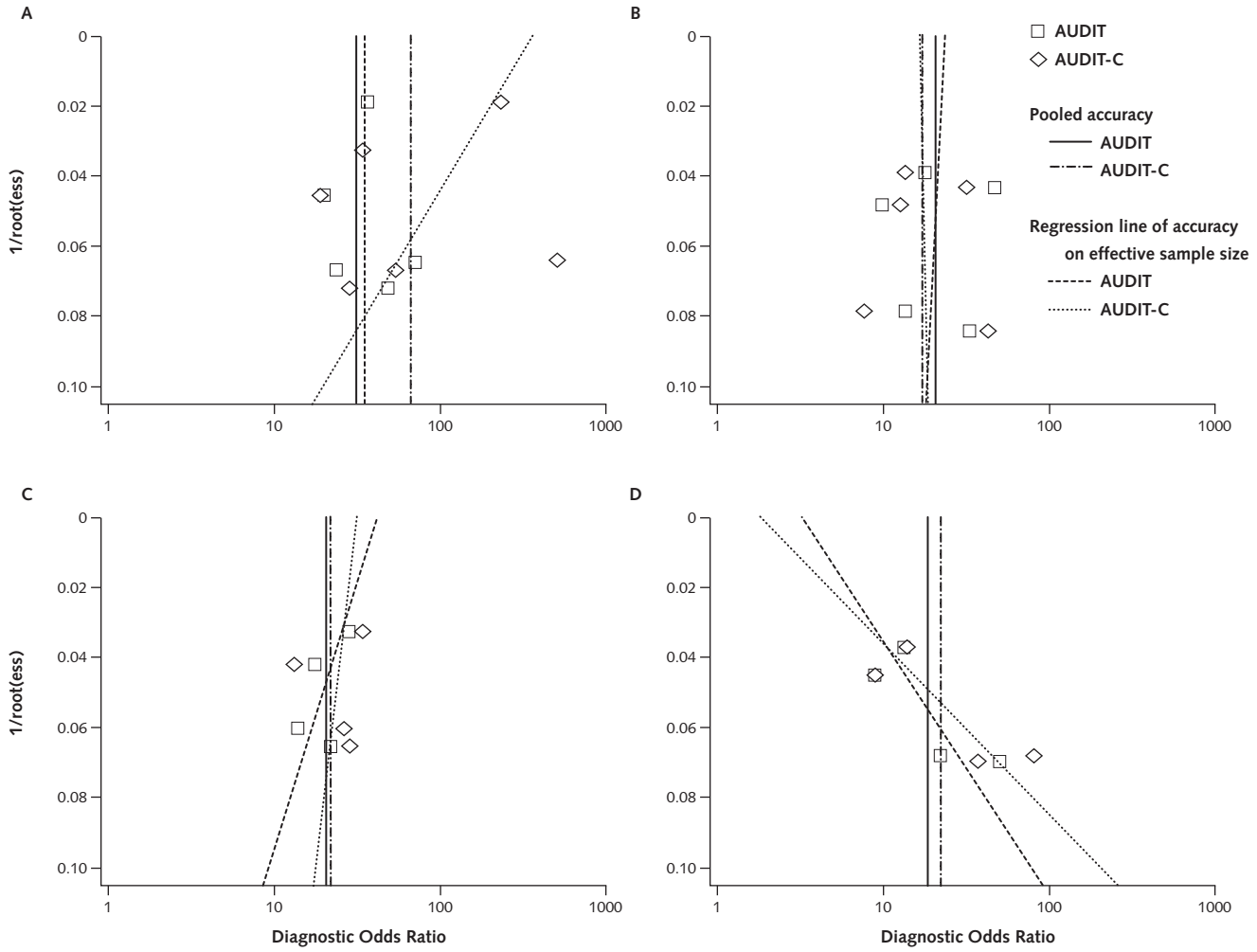
Solid circles indicate summary point estimates, dotted lines indicate 95% CIs, and pairs of connected points indicate single studies. **A.** Accuracy for risky drinking in primary care. **B.** Accuracy for alcohol use disorders in primary care. **C.** Accuracy for unhealthy alcohol use in primary care. **D.** Accuracy for risky drinking in the general population.

Appendix Table 4. **Regression of Accuracy against Effective Sample Size to Test for Possible Small-Study Effects**

Target Condition	Standardized Regression Coefficient	Level of Significance for Testing whether Coefficient Differs from Zero (*P* Value)
Risky drinking in primary care*		
AUDIT	−0.005	0.99
AUDIT-C	−0.564	0.32
Alcohol use disorder in primary care*		
AUDIT	−0.057	0.93
AUDIT-C	0.032	0.96
Unhealthy alcohol use in primary care†		
AUDIT	−0.674	0.33
AUDIT-C	−0.172	0.83
Risky drinking in general population†		
AUDIT	0.767	0.23
AUDIT-C	0.829	0.171

AUDIT = Alcohol Use Disorders Identification Test; AUDIT-C = Alcohol Use Disorders Identification Test–Consumption.
* Based on 5 studies.
† Based on 4 studies.

Appendix Figure 3. **Funnel plots of accuracy against effective sample size to test for possible small-study effects.**



AUDIT = Alcohol Use Disorders Identification Test; AUDIT-C = Alcohol Use Disorders Identification Test–Consumption; ess = effective sample size. **A.** Risky drinking in primary care. **B.** Alcohol use disorders in primary care. **C.** Unhealthy alcohol use in primary care. **D.** Risky drinking in the general population.