Changes in alcohol consumption reported on routine healthcare screenings are associated with changes in depression symptoms

Kevin A. Hallgren, PhD[1,2,3], **Helen E. Jack, MD**[4], **Malia Oliver, BA**[2], **Douglas Berger, MD MLitt**[4,5], **Jennifer F. Bobb, PhD**[2], **Daniel R. Kivlahan, PhD**[6], **Katharine A. Bradley, MD MPH**[2,3,4]

[1]Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA, United States

[2]Kaiser Permanente Washington Health Research Institute, Seattle, WA, United States

[3]University of Washington, Department of Health Systems and Population Health, Seattle, WA, United States

[4]Division of General Internal Medicine, University of Washington, Seattle, United States

[5]General Medicine Service VA Puget Sound, Seattle, WA, United States

[6]Center of Innovation for Veteran-Centered and Value-Driven Care, Health Services Research and Development, Veteran Affairs Puget Sound HealthCare System, Seattle, WA, United States

Abstract

Background: The Alcohol Use Disorders Identification Test-Consumption version (AUDIT-C) has been robustly validated as a point-in-time screen for unhealthy alcohol use, but less is known about the significance of changes in AUDIT-C scores from routine screenings over time. Unhealthy alcohol use and depression commonly co-occur, and changes in drinking often co-occur with changes in depression symptoms. We assess the associations of changes in AUDIT-C scores and changes in depression symptoms reported on brief screens completed in routine care.

Methods: The study sample included 198,335 primary care patients who completed two AUDIT-C screens 11–24 months apart and the Patient Health Questionnaire-2 (PHQ-2) depression screen on the same day as each AUDIT-C. Both screening measures were completed as part of routine care within a large health system in Washington state. AUDIT-C scores were categorized to reflect 5 drinking levels at both time points, resulting in 25 subgroups with different change patterns. For each of the 25 subgroups, within-group changes in the prevalence of positive PHQ-2 depression screens were characterized using risk ratios (RRs) and McNemar's tests.

Results: Patient subgroups with increases in AUDIT-C risk categories generally experienced increases in the prevalence of positive depression screens (RRs ranging from 0.95 to 2.00). Patient subgroups with decreases in AUDIT-C risk categories generally experienced decreases in the

Correspondence: Kevin A. Hallgren, PhD. Address: Department of Psychiatry and Behavioral Sciences, University of Washington Medical Center, Box 356560, 1959 NE Pacific St., Seattle, WA 98195, United States. khallgre@uw.edu. Phone: +1 (206) 616-2906. Fax: +1 (206) 543-9520.

prevalence of positive depression screens (RRs ranging from 0.52 to 1.01). Patient subgroups that did not have changes in AUDIT-C risk categories experienced little or no change in prevalence of positive depression screens (RRs ranging from 0.98 to 1.15).

Conclusions: Changes in alcohol consumption reported on AUDIT-C screens completed in routine care were associated with changes in depression screening results in the hypothesized manner. Results support the validity and clinical utility of monitoring changes in AUDIT-C scores over time as a meaningful measure of changes in drinking.

Keywords

alcohol use; screening; changes in drinking; depression; measurement-based care; mental health; primary care

INTRODUCTION

Unhealthy alcohol use reflects a spectrum of alcohol use that ranges from drinking above recommended limits up to meeting criteria for alcohol use disorder (AUD) (Saitz, 2005). Unhealthy alcohol use is associated with increased risk for mental health problems, including depression symptoms (Grant et al., 2004; McHugh & Weiss, 2019; Ryan et al., 2022), emergency department encounters (Suen et al., 2022), and hospitalizations for mental health conditions (Davis et al., 2010; Fortney et al., 1999). Epidemiological and treatment studies also show that abstinence or decreased drinking in people with AUD is associated with decreases in depression symptoms (Hallgren et al., 2018; Hallgren, Holzhauer, et al., 2021; Liappas et al., 2002).

In primary care settings, up to 20–30% of patients may have unhealthy alcohol use and could benefit from brief interventions focused on reducing alcohol consumption (Bradley et al., 2007; Saitz, 2005). Primary care settings are increasingly implementing annual screenings for unhealthy alcohol use and depression symptoms as part of standard care due to the high prevalence of these conditions in adult populations, the high degree of morbidity associated with them, and the ability to address both issues in primary care settings (Donroe & Edelman, 2022). The US Preventive Services Task Force (Siu et al., 2016; US Preventive Services Task Force et al., 2018) also recommends such screenings to help providers identify and provide follow-up care for depression and unhealthy alcohol use.

To further support alcohol-related clinical care within primary care, there is a need for screening tools that are valid, practical to administer in routine care, and can be used to meaningfully monitor changes in drinking over time. The Alcohol Use Disorders Identification Test-Consumption version (AUDIT-C) is a validated measure of alcohol consumption (Bradley et al., 2007; Rubinsky et al., 2013) that is freely available and practical to administer in many primary care settings to identify people who drink above recommended limits. However, limited research has been conducted to evaluate AUDIT-C scores as a longitudinal measure of changes in drinking over time, including whether changes in AUDIT-C are associated with changes in other meaningful clinical outcomes.

Previous work in Veteran's Health Administration (VA) samples has shown that changes in AUDIT-C scores completed in routine care approximately 1–2 years apart are associated with high-density lipoprotein cholesterol, the incidence of alcohol-related gastrointestinal hospitalizations, and the incidence of physical trauma in the general VA population (Bradley et al., 2016; Rubinsky et al., 2019) as well as changes in HIV disease severity and HIV medication adherence in VA patients with HIV (Williams et al., 2018, 2019, 2021). However, to our knowledge, no studies have evaluated longitudinal changes in AUDIT-C scores completed by primary care patients in routine care outside of a VA setting. This limits generalizability considering that the VA samples were predominantly composed of men (e.g., 94–97% male) and often older adults (e.g., 57–60% over age 65 in non-HIV studies), especially since drinking patterns and the prevalence of mental health problems can differ in samples composed of younger, non-veteran, and more often female individuals (Fuehrlein et al., 2016). Additionally, we are unaware of studies that have evaluated whether changes in AUDIT-C scores completed by primary care patients as part of routine clinical screening are associated with changes in mental health measures, even though such associations have been found in epidemiological and treatment studies (Hallgren et al., 2018; Hallgren, Holzhauer, et al., 2021; Knox et al., 2019). Understanding the relationship between changes in AUDIT-C scores and changes in mental health symptoms may help clarify the clinical utility of the AUDIT-C as a longitudinal measure. It could inform the designs of pragmatic trials that may monitor changes in AUDIT-C scores as clinical outcomes and may help clarify the predictive validity of changes in AUDIT-C scores in relation to depression symptoms. Such work would extend previous studies that have demonstrated the clinical utility and validity of AUDIT-C's completed as a single point-in-time measure in routine care settings (Bradley et al., 2003, 2007, 2016; Rubinsky et al., 2013; Ryan et al., 2022) by highlighting the potential utility of using AUDIT-C scores to monitor longitudinal changes in drinking in routine care settings.

The aim of the current study was to evaluate whether changes in drinking measured by AUDIT-C screens completed in routine care 11–24 months apart were associated with changes in the prevalence of positive depression screens completed on the same day as the alcohol screenings. We hypothesized that (1) patients with increases in AUDIT-C risk score categories over an 11–24 month period would tend to have increases in the prevalence of positive depression screens, (2) patients with decreases in AUDIT-C risk score categories over the same timeframe would tend to have decreases in the prevalence of positive depression screens, and (3) patients with no change in AUDIT-C risk score categories would tend to have no change in the prevalence of positive depression screens.

MATERIALS AND METHODS

Study Setting and Population

This is a retrospective cohort study using clinical data from electronic health records (EHRs) and administrative claims from the Kaiser Permanente (KP) Washington health system. KP Washington is an integrated healthcare system in Washington state that provides both health insurance and integrated primary and specialty care. The study was approved by the KP Washington Health Research Institute's Institutional Review Board with waivers of consent

and Health Insurance Portability and Accountability Act authorization to use existing EHR data.

KP Washington currently has 30 primary care practices that have implemented annual alcohol screening via the AUDIT-C and depression screening via the Patient Health Questionnaire-2 (PHQ-2) (Kroenke et al., 2003) to support behavioral health integration across the health system (Glass et al., 2018; Lee et al., In Press; Yeung et al., 2020). Most screens are completed in primary care settings where the procedures for completing them are standardized (i.e., the EHR prompts medical assistants to administer screens on paper forms, which patients complete in waiting areas or exam rooms and medical assistants enter responses into the EHR); however, screens may also be completed in specialty clinics. Since summer 2020, much of the behavioral health screening has been completed via the online EHR portal prior to visits or on tablets in waiting rooms.

Patients were included in this study if they (a) had 1 visit to a KP Washington primary care practice during the study observation period of October 1, 2016 to October 31, 2020, (b) completed two AUDIT-C screens 11–24 months apart during the study observation period, (c) completed PHQ-2 depression screens on the same day as both AUDIT-C screens, and (d) were at least 18 years old when both AUDIT-C screens were completed. If a patient had more than one eligible pair of AUDIT-C screens within the study window, a random AUDIT-C was selected as the first screen (T1) and the closest following AUDIT-C that was completed 11–24 months later was selected as the second screen (T2). We selected an 11–24 month timeframe between the two AUDIT-C screens because it reflected the annual screening window used within the health system (i.e., screens usually administered at clinical encounters if they have not been completed within the past year, although sometimes screens were administered a few days or weeks before they were due). The timeframe also accounted for the fact that many patients utilize primary care less than once per year and minimized potential overlap between the two AUDIT-C lookback periods (i.e., the AUDIT-C instructs patients to report on their drinking over the past year).

Measures

Alcohol Screening Measure—The AUDIT-C (Bradley et al., 2003; Bush et al., 1998) consists of the first three items of the World Health Organization AUDIT measure, which ask patients to report their frequency of drinking, typical drinks per drinking day, and frequency of drinking 6 or more drinks in a day over the past year. Each item is answered on a 5-point scale (0–4 points). Responses are summed to derive a total score (0–12 points). For the current study, we derived 5 drinking levels (see Table 1) using AUDIT-C scores that reflect clinically meaningful categories based on thresholds that reflect any drinking (scores 1), unhealthy alcohol use that typically warrants brief intervention in primary care (scores 3 female, 4 male) (Bradley et al., 2007), high-risk drinking that is typically followed by assessment of AUD symptoms in the KP Washington health system (scores 7) (Sayre et al., 2020), and scores that reflect exponentially increasing levels of alcohol consumption and marked increase in risk for AUD a(Au et al., 2007; Bradley et al., 2016; Jack et al., In Press; Rubinsky et al., 2012, 2013; Sayre et al., 2020; Williams et al., 2012)nd adverse health outcomes (scores 9) (Au et al., 2007; Bradley et al., 2016; Jack et al., In Press; Rubinsky

et al., 2012, 2013; Williams et al., 2012). By consolidating the 13-point scale (scores 0–12) into 5 drinking risk categories, we reduced the number of combinations of drinking risk levels at T1 and T2 (i.e., 5×5 = 25 subgroups instead of 13×13 = 169 subgroups), making statistical analyses more parsimonious and increasing precision of estimates within subgroups. Although AUDIT-C scores are often classified into just 2 categories in some research and clinical contexts (i.e., presence or absence of unhealthy alcohol use based on scores   3 for women or   4 for men), the use of 5 categories for the current study allowed for examinations of scaled reductions in drinking, including in subgroups that potentially reported increases or decreases in drinking risk levels yet remained above the threshold for unhealthy alcohol use at both time points.

Depression Screening Measure—The PHQ-2 is a two-item validated depression screening questionnaire (Arroll et al., 2010; Kroenke et al., 2003; Löwe et al., 2005). At KP Washington, the PHQ-2 is included on the same annual behavioral health screen as the AUDIT-C. The two items ask patients to report the frequency of depressed mood and anhedonia over the past two weeks, with response options ranging from 0 ("not at all") to 3 ("nearly every day"). Responses to both items are summed to yield total scores ranging from 0–6, which are used to create a dichotomous outcome reflecting screens that are "positive" and indicate a high likelihood of a current depression disorder (PHQ-2 scores 3–6) or screens that are "negative" indicating a low likelihood of a current depression disorder (scores 0–2). This cutoff score of 3 was selected because it has moderate sensitivity and high specificity for major depressive disorder (Arroll et al., 2010) and is often used as a cutoff score in clinical settings for triggering more in-depth depression-related assessment and care. Previous research has indicated that changes in PHQ-2 scores are sensitive to detecting within-group changes in symptoms of depressive disorders (Löwe et al., 2005). Although depression symptoms exist along a continuum of severity and longer instruments can capture the severity of depression symptoms (e.g., PHQ-9, Kroenke et al., 2001), longer measures were not available for most patients in this sample.

Descriptive measures—Age, sex, race, and ethnicity data were obtained from the KP Washington EHR. Health insurance plan payer (e.g., Medicare, Medicaid, employer-paid, self-pay) was obtained from KP Washington records. Outpatient AUD and mental health diagnoses up to two years prior to the first AUDIT-C were obtained from ICD-10 codes documented in the KP Washington EHR and insurance claims while patients were enrolled in a KP Washington health plan.

Analytic Approach

Descriptive analyses characterized patients' demographics, past two-year AUD diagnoses, substance use disorder diagnoses, mental health disorder diagnoses, and AUDIT-C scores. Primary statistical analyses examined whether patient subgroups experienced significant changes in the prevalence of positive depression screens from the time of the first AUDIT-C (T1) to the time of the second AUDIT-C (T2). Tests for within-group changes in these outcomes were performed using McNemar's tests, which are non-parametric tests that determine whether there are significant changes in the proportions of a binary outcome over matched pairs of observations (e.g., significant within-group changes in the proportion

of positive depression screens at T1 versus T2). These within-group analyses provided the benefit of modeling changes in the prevalence of positive depression screens over time in a manner that allows each patient to serve as their own comparator. This is particularly useful in the current study, as many variables that are related to mental health outcomes (e.g., sex, age, race, ethnicity, socioeconomic status, genetic factors) remain relatively stable within persons over an 11–24-month period. Allowing each person to serve as their own comparator also increases statistical power for detecting changes in the outcome measures relative to between-group tests. A series of McNemar's tests was conducted within stratified subgroups reflecting how drinking (i.e., the 5 AUDIT-C score categories) can change from T1 to T2 (i.e., 5×5 total combinations). This stratification allowed us to test for within-group changes for each subgroup reflecting a unique pattern of changes in AUDIT-C score categories. This stratification also allowed us to avoid the assumption of linear associations between the AUDIT-C score categories and the prevalence of positive depression screens, which often does not hold (e.g., Bradley et al., 2016; Ryan et al., 2022). Due to the large number of subgroups reflecting different patterns of changes in AUDIT-C score categories, we took a descriptive approach to interpreting the results by focusing on patterns of significance and effect sizes reflecting within-group change across the 5×5 subgroups, rather than focusing our interpretation on each subgroup individually (Bradley et al., 2016). To facilitate interpretation of effect sizes for changes in within-group risk, we computed risk ratios (RR) that reflected the magnitude of change in the prevalence of positive depression screens from T1 to T2 within each subgroup (i.e., prevalence of positive depression screens at T2 divided by the prevalence of positive depression screens at T1). Precision of the RR estimates was characterized by estimating 95% CI's via nonparametric bootstrapping that resampled patients 10,000 times within each subgroup.

RESULTS

Description of sample

There were 198,335 patients who met study inclusion criteria. Table 2 displays their demographics, substance use and mental health disorder diagnoses, and AUDIT-C scores. Most patients were female, white, and commercially insured (i.e., insurance paid for by an employer or self-pay); the largest age group was 45–64. On average, patients completed the T1 and T2 AUDIT-C's 16.22 months apart (SD=3.62). Compared to adult primary care patients who completed alcohol and depression screens but *didn't* have two screens 11–24 months apart (i.e., those *not* included in the current sample due to a lack of longitudinal screens), the patients included in the current sample were more often age 45+, female, white, insured by Medicaid or Medicare, and more likely to have documented mental health, alcohol, and other substance use disorder diagnoses within the past two years (see eSupplement).

Changes in AUDIT-C scores and changes in depression screens.

The prevalence of positive depression screens was 13.4% at T1 and 13.5% at T2. The prevalence of AUDIT-C scores reflecting unhealthy alcohol use (scores 3–12 for women, 4–12 for men) were 26.6% at T1 and 25.3% at T2. The 5×5 grid in Figure 1 shows within-group changes in the prevalence of positive depression screens for each patient

subgroup with different combinations of AUDIT-C score categories at T1 (rows in grid) and T2 (columns in grid). For each graph, the heights of the two bars reflect the prevalence of positive depression screens at T1 and T2 within each subgroup; whiskers reflect the 95% CI's of these prevalence estimates. Table 3 displays the effect size estimates (RR's and 95% CI's) reflecting changes in the prevalence of positive depression screens from T1 to T2 within all 5×5 subgroups. The point estimates and 95% CI's depicted in Figure 1 are available in the eSupplement.

Patients who had no changes in AUDIT-C score categories from T1 to T2 are presented in the gray graphs on the left-to-right diagonal of the 5×5 grid in Figure 1. In total, 73.7% of the sample had the same AUDIT-C risk category at both time points. The prevalence of positive depression screens did not change significantly for patient subgroups who, at both time points, remained in the same AUDIT-C score categories reflecting no drinking (AUDIT-C score 0) or unhealthy alcohol use with high-risk drinking (AUDIT-C scores 7–8) or very high-risk drinking (AUDIT-C scores 9–12). The prevalence of positive depression screens significantly increased for patient subgroups who, at both time points, remained in the same AUDIT-C score categories reflecting drinking without unhealthy alcohol use (AUDIT-C scores 1–2 female, 1–3 male) or unhealthy alcohol use with moderate-risk drinking (AUDIT-C scores 3–6 female, 4–6 male). However, the magnitudes of these differences were small, RR=1.03 (95% CI: 1.00–1.05) and 1.05 (95% CI: 1.01–1.09), respectively, reflecting 3% and 5% relative increases in the prevalence of positive depression screens at T2 over T1 (see Table 3).

Patients who transitioned from a lower AUDIT-C score category at T1 to a higher category at T2 are displayed in the orange graphs above the diagonal of the 5×5 grid in Figure 1. For each subgroup with an increase in drinking risk level, there was a significant increase in the prevalence of positive depression screens except within the subgroup who increased their AUDIT-C scores from no drinking at T1 (AUDIT-C score 0) to drinking without unhealthy alcohol use at T2 (AUDIT-C scores 1–2 female, 1–3 male) and the subgroup who had unhealthy alcohol use with high-risk drinking at T1 (AUDIT-C scores 7–8) and unhealthy alcohol use with very high-risk drinking at T2 (AUDIT-C scores 9–12). For all other subgroups reflecting increased AUDIT-C scores, there were significant increases in the prevalence of positive depression screens, with the prevalence of positive depression screens ranging from 11% to 100% higher at T2 compared to T1 (see Table 3 for RR's and 95% CI's for all subgroups).

Patients who transitioned from a higher AUDIT-C score category at T1 to a lower category at T2 are displayed in the blue graphs below the left to right diagonal of the 5×5 grid in Figure 1. For each subgroup with a decrease in drinking risk level, there was a significant reduction in the prevalence of positive depression screens, except within the subgroup of patients who reported drinking without unhealthy alcohol use at T1 (AUDIT-C scores 1–2 female, 1–3 male) then reported no drinking at T2 (AUDIT-C score 0). For all other subgroups reflecting reduced AUDIT-C scores, reductions in the prevalence of positive depression screens were significant, with the prevalence of positive depression screens ranging from 17% to 49% lower at T2 compared to T1 (see Table 3 for RR's and 95% CI's for all subgroups).

DISCUSSION

The current study found that changes in scores on the AUDIT-C completed by primary care patients in routine care 11–24 months apart were generally associated with changes in the prevalence of positive depression screens over the same period. Patients with increases in AUDIT-C scores tended to have increases in the prevalence of positive depression screens, patients with reductions in AUDIT-C scores tended to have decreases in the prevalence of positive depression screens, and patients with little or no changes in AUDIT-C scores tended to have little or no changes in the prevalence of positive depression screens. To our awareness, this is the first study to evaluate the associations of changes in AUDIT-C scores completed in routine care and changes in mental health outcomes. It is also the first study to our awareness that evaluates changes in routine care AUDIT-C scores and changes in any health outcome measured for patients outside of the VA.

The results of this study have important implications for clinicians. The observed longitudinal associations between changes in AUDIT-C scores and changes in PHQ-2 depression screening results support the validity and clinical utility of monitoring changes in AUDIT-C scores over time as a meaningful measure of changes in drinking. Annual screening with the AUDIT-C may allow clinicians to not only monitor current drinking risk and identify patients who could potentially benefit from brief interventions, but also to assess *changes in* AUDIT-C scores over time to understand changes in drinking. For example, clinicians could monitor changes in AUDIT-C scores to identify patients who have reduced or increased their drinking, which could inform their assessments and clinical decision making related to alcohol use.

While many clinicians may be accustomed to thinking about alcohol screening results in binary terms that indicate the presence or absence of unhealthy alcohol use, changes in AUDIT-C scores can provide additional information and could prompt useful conversations even when patients remain above or below a screening threshold. For example, the results of the current study suggest that patients who consistently screen above the threshold for unhealthy alcohol use, but still decreased their drinking, had significant reductions in risk for depression even with continued unhealthy alcohol use. Likewise, patients who reported non-abstinent drinking reductions also had reductions in depression symptoms, including patients who had AUDIT-C scores at both time points that were high enough to suggest a high risk for AUD (AUDIT-C 7)(Hallgren, Matson, et al., 2021; Sayre et al., 2020). This finding contributes to a larger body of literature emphasizing reductions in drinking risk levels (Knox et al., 2019; Witkiewitz et al., 2017, 2020) as reflecting clinically meaningful outcomes and suggests that clinicians should focus on interpreting changes in AUDIT-C scores on a scaled continuum, rather taking a purely dichotomous interpretation that focuses only on whether the AUDIT-C score is above or below a given threshold. The current study created 5 drinking risk level categories to facilitate more interpretable subgroup analyses, and clinicians may choose to utilize these categories when interpreting AUDIT-C scores or they may find it heuristically simpler to forgo categorization and interpret their patients' AUDIT-C scores along a scaled continuum of drinking risk.

The results of this study also have important implications for patients. Our results are consistent with prior research conducted outside of routine care settings showing an association between changes in drinking and changes in symptoms of depression (Hallgren, Holzhauer, et al., 2021; Knox et al., 2019; Liappas et al., 2002). Importantly, the results of the current study should not be interpreted as suggesting a specific type of causal relationship between drinking and depression. The results observed here could be attributed to changes in drinking causing changes in depression, changes in depression causing changes in drinking, or changes in other variables causing changes in both depression and drinking. Nonetheless, for patients who have both depression symptoms and unhealthy alcohol use, learning that reductions in drinking are often associated with concurrent reductions in depression symptoms may help foster motivation for addressing one or both conditions. Primary care providers could use these results to engage patients in discussions about the complex relationship between alcohol use and depression and consideration of treatment options for both drinking and depression (Ramsey et al., 2005).

The current study also has important implications for alcohol research. Researchers conducting population-based studies, pragmatic trials, and implementation studies often need valid and easy-to-use measures that are documented as part of routine care, like the AUDIT-C, to measure changes in drinking and related risk factors in healthcare systems. Findings of the current study and previous studies in VA settings (Bradley et al., 2016; Rubinsky et al., 2019; Williams et al., 2018, 2019) show that population-level changes in AUDIT-C scores are associated with meaningful changes in health outcomes. Collectively, these studies provide evidence that changes in AUDIT-C scores may reflect changes in drinking, and the current study suggests they reflect corresponding changes in risk for depression symptoms.

Our study has noteworthy limitations. Because we evaluated alcohol and depression screening measures that were completed in routine care, we were unable to measure or control for many factors that could impact patient self-reports on these measures (e.g., patients' experiences with medical assistants administering the screens, social desirability, stigma). As the alcohol and depression screens were administered together, factors influencing self-report on a given day might impact both scores, and unmodeled time-varying confounders (e.g., changes in employment, health status, etc.) could not be accounted for. Future research should therefore evaluate associations of changes in AUDIT-C scores with changes in longer-range outcomes that do not rely on self-report, including use of emergency and acute mental health care services. It is also possible that administering the AUDIT-C as part of routine care, entering the results into EHRs, and knowing that results will be viewed by healthcare providers led many patients to underreport their drinking or depression symptoms. The two-item depression screen was utilized as a dichotomous measure based on previous validation studies and therefore did not differentiate the severity of depression symptoms. The study did not attempt to identify causal relationships between drinking and depression and focused specifically on within-group changes, and therefore it did not control for factors that could differ between subgroups that could impact changes in drinking and mental health. While our sample demographics were approximately consistent with the overall population of Washington State, it was predominantly white and non-Hispanic and mostly had commercial insurance

or Medicare; thus additional research focusing on more marginalized groups is warranted, including with people of color, people with lower socioeconomic status, and those who have less access to primary healthcare services. Patients who had two AUDIT-C scores during the study period (and thus were included in the sample) were older and more likely to be commercially insured and have mental health and/or substance use disorders within the past two years compared to those who had only a single AUDIT-C during the observation window (and were thus excluded from the study); this may reflect differences in healthcare utilization, where individuals with greater medical or psychiatric needs have had more contact with the health system, which in turn increased their opportunity for repeated screenings during the study period. Therefore, study sample may have included a population with higher risk for unhealthy alcohol use and depression than the general primary care population. Studies using a longer time-window than 24 months would potentially include more patients who receive care infrequently and assess the value of changes in AUDIT-C over a longer period.

Our study also had several strengths. The AUDIT-C and PHQ-2 measures were obtained as part of routine-care, and thus our study provides evidence of how these measures are likely to perform under real-world conditions. This enhances external validity, as the results reflect how patients respond on alcohol and depression screens that are completed in healthcare settings (i.e., where information about drinking – a behavior that is frequently stigmatized – is documented in patients' medical records and seen by patients' medical providers), rather than how participants would respond in a research setting (i.e., where information about drinking typically is not entered into one's medical record or seen by their medical providers). The screening measures were also brief, validated, and practical for use in routine care, and the results therefore complement previous epidemiological and treatment studies that have shown similar relationships with longer measures that often use optimal assessment techniques (e.g., structured interviews) may not be practical for use in routine care. The use of a routine care sample allowed us to obtain a large sample size and likely limited the risk of sampling bias that would have occurred if we only included people who were willing to participate in alcohol-related research. The use of a within-group statistical model allowed each patient to serve as their own comparison, which increased power and reduced the influence of stable potential confounding variables.

Conclusion

Increases and decreases in AUDIT-C scores completed in routine care 11–24 months apart are associated with increases and decreases, respectively, in the prevalence of positive depression screens. In addition to being valid, reliable, and practical to administer for point-in-time identification of unhealthy alcohol use, the AUDIT-C may also be useful as a longitudinal, scaled measure that clinicians and researchers can utilize to monitor changes in drinking.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

References

Arroll B, Goodyear-Smith F, Crengle S, Gunn J, Kerse N, Fishman T, Falloon K, & Hatcher S (2010). Validation of PHQ-2 and PHQ-9 to Screen for Major Depression in the Primary Care Population. Annals of Family Medicine, 8(4), 348–353. 10.1370/afm.1139 [PubMed: 20644190]

Au DH, Kivlahan DR, Bryson CL, Blough D, & Bradley KA (2007). Alcohol screening scores and risk of hospitalizations for GI conditions in men. Alcoholism, Clinical and Experimental Research, 31(3), 443–451. 10.1111/j.1530-0277.2006.00325.x [PubMed: 17295729]

Bradley KA, Bush KR, Epler AJ, Dobie DJ, Davis TM, Sporleder JL, Maynard C, Burman ML, & Kivlahan DR (2003). Two brief alcohol-screening tests From the Alcohol Use Disorders Identification Test (AUDIT): Validation in a female Veterans Affairs patient population. Archives of Internal Medicine, 163(7), 821–829. 10.1001/archinte.163.7.821 [PubMed: 12695273]

Bradley KA, DeBenedetti AF, Volk RJ, Williams EC, Frank D, & Kivlahan DR (2007). AUDIT-C as a brief screen for alcohol misuse in primary care. Alcoholism, Clinical and Experimental Research, 31(7), 1208–1217. 10.1111/j.1530-0277.2007.00403.x [PubMed: 17451397]

Bradley KA, Rubinsky AD, Lapham GT, Berger D, Bryson C, Achtmeyer C, Hawkins EJ, Chavez LJ, Williams EC, & Kivlahan DR (2016). Predictive validity of clinical AUDIT-C alcohol screening scores and changes in scores for three objective alcohol-related outcomes in a Veterans Affairs population. Addiction (Abingdon, England), 111(11), 1975–1984. 10.1111/add.13505 [PubMed: 27349855]

Bush K, Kivlahan DR, McDonell MB, Fihn SD, & Bradley KA (1998). 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. Archives of Internal Medicine, 158(16), 1789–1795. 10.1001/archinte.158.16.1789 [PubMed: 9738608]

Davis LL, Wisniewski SR, Howland RH, Trivedi MH, Husain MM, Fava M, McGrath PJ, Balasubramani GK, Warden D, & Rush AJ (2010). Does comorbid substance use disorder impair recovery from major depression with SSRI treatment? An analysis of the STAR*D level one treatment outcomes. Drug and Alcohol Dependence, 107(2), 161–170. 10.1016/j.drugalcdep.2009.10.003 [PubMed: 19945804]

Donroe JH, & Edelman EJ (2022). Alcohol Use. Annals of Internal Medicine, 175(10), ITC145–ITC160. 10.7326/AITC202210180 [PubMed: 36215713]

Fortney JC, Booth BM, & Curran GM (1999). Do Patients with Alcohol Dependence Use More Services? A Comparative Analysis with other Chronic Disorders. Alcoholism: Clinical and Experimental Research, 23(1), 127–133. 10.1111/j.1530-0277.1999.tb04033.x [PubMed: 10029213]

Fuehrlein BS, Mota N, Arias AJ, Trevisan LA, Kachadourian LK, Krystal JH, Southwick SM, & Pietrzak RH (2016). The burden of alcohol use disorders in US military veterans: Results from the National Health and Resilience in Veterans Study. Addiction (Abingdon, England), 111(10), 1786–1794. 10.1111/add.13423 [PubMed: 27061707]

Glass JE, Bobb JF, Lee AK, Richards JE, Lapham GT, Ludman E, Achtmeyer C, Caldeiro RM, Parrish R, Williams EC, Lozano P, & Bradley KA (2018). Study protocol: A cluster-randomized

trial implementing Sustained Patient-centered Alcohol-related Care (SPARC trial). Implementation Science: IS, 13(1), 108. 10.1186/s13012-018-0795-9 [PubMed: 30081930]

Grant BF, Stinson FS, Dawson DA, Chou SP, Dufour MC, Compton W, Pickering RP, & Kaplan K (2004). Prevalence and Co-occurrence of Substance Use Disorders and IndependentMood and Anxiety Disorders: Results From the National Epidemiologic Survey on Alcohol and RelatedConditions. Archives of General Psychiatry, 61(8), 807–816. 10.1001/archpsyc.61.8.807 [PubMed: 15289279]

Hallgren KA, Delker BC, & Simpson TL (2018). Effects of initiating abstinence from alcohol on daily craving and negative affect: Results from a pharmacotherapy clinical trial. Alcoholism: Clinical and Experimental Research, 42(3), 634–645. [PubMed: 29286542]

Hallgren KA, Holzhauer CG, Epstein EE, McCrady BS, & Cook S (2021). Optimizing the length and reliability of measures of mechanisms of change to support measurement-based care in alcohol use disorder treatment. Journal of Consulting and Clinical Psychology, 89(4), 277–287. 10.1037/ ccp0000643 [PubMed: 34014690]

Hallgren KA, Matson TE, Oliver M, Witkiewitz K, Bobb JF, Lee AK, Caldeiro RM, Kivlahan D, & Bradley KA (2021). Practical Assessment of Alcohol Use Disorder in Routine Primary Care: Performance of an Alcohol Symptom Checklist. Journal of General Internal Medicine 10.1007/ s11606-021-07038-3

Jack H, Berger D, Kivlahan DR, Bradley KA, & Hallgren KA (In Press). Association between clinical measures of unhealthy alcohol use and subsequent year hospital admissions in a primary care population. Drug and Alcohol Dependence

Knox J, Scodes J, Wall M, Witkiewitz K, Kranzler HR, Falk D, Litten R, Mann K, O'Malley SS, Anton R, & Hasin DS (2019). Reduction in non-abstinent WHO drinking risk levels and depression/anxiety disorders: 3-year follow-up results in the US general population. Drug and Alcohol Dependence, 197, 228–235. 10.1016/j.drugalcdep.2019.01.009 [PubMed: 30852375]

Kroenke K, Spitzer RL, & Williams JB (2001). The PHQ-9: Validity of a brief depression severity measure. Journal of General Internal Medicine, 16(9), 606–613. 10.1046/ j.1525-1497.2001.016009606.x [PubMed: 11556941]

Kroenke K, Spitzer RL, & Williams JBW (2003). The Patient Health Questionnaire-2: Validity of a two-item depression screener. Medical Care, 41(11), 1284–1292. 10.1097/01.MLR.0000093487.78664.3C [PubMed: 14583691]

Lee AK, Bobb JF, Richards JE, Achtmeyer CE, Ludman EJ, Oliver M, Caldeiro RM, Parrish R, Lozano P, Lapham GT, Williams EC, Glass JE, & Bradley KA (In Press). Integrating alcohol-related care into primary care as part of behavioral health integration: A cluster-randomized implementation trial. JAMA Internal Medicine

Liappas J, Paparrigopoulos T, Tzavellas E, & Christodoulou G (2002). Impact of alcohol detoxification on anxiety and depressive symptoms. Drug and Alcohol Dependence, 68(2), 215–220. 10.1016/ s0376-8716(02)00195-3 [PubMed: 12234651]

Löwe B, Kroenke K, & Gräfe K (2005). Detecting and monitoring depression with a two-item questionnaire (PHQ-2). Journal of Psychosomatic Research, 58(2), 163–171. 10.1016/ j.jpsychores.2004.09.006 [PubMed: 15820844]

McHugh RK, & Weiss RD (2019). Alcohol Use Disorder and Depressive Disorders. Alcohol Research: Current Reviews, 40(1), arcr.v40.1.01. 10.35946/arcr.v40.1.01

Ramsey SE, Engler PA, & Stein MD (2005). Alcohol Use Among Depressed Patients: The Need for Assessment and Intervention. Professional Psychology, Research and Practice, 36(2), 203–207. 10.1037/0735-7028.36.2.203 [PubMed: 20502623]

Rubinsky AD, Chavez LJ, Berger D, Lapham GT, Hawkins EJ, Williams EC, & Bradley KA (2019). Utility of routine alcohol screening for monitoring changes in alcohol consumption. Drug and Alcohol Dependence, 201, 155–160. 10.1016/j.drugalcdep.2019.03.010 [PubMed: 31229703]

Rubinsky AD, Dawson DA, Williams EC, Kivlahan DR, & Bradley KA (2013). AUDIT-C scores as a scaled marker of mean daily drinking, alcohol use disorder severity, and probability of alcohol dependence in a U.S. general population sample of drinkers. Alcoholism, Clinical and Experimental Research, 37(8), 1380–1390. 10.1111/acer.12092 [PubMed: 23906469]

Rubinsky AD, Sun H, Blough DK, Maynard C, Bryson CL, Harris AH, Hawkins EJ, Beste LA, Henderson WG, Hawn MT, Hughes G, Bishop MJ, Etzioni R, Tønnesen H, Kivlahan DR, & Bradley KA (2012). AUDIT-C alcohol screening results and postoperative inpatient health care use. Journal of the American College of Surgeons, 214(3), 296–305.e1. 10.1016/j.jamcollsurg.2011.11.007 [PubMed: 22244208]

Ryan ED, Chang YM, Oliver M, Bradley KA, & Hallgren KA (2022). An Alcohol Symptom Checklist identifies high rates of alcohol use disorder in primary care patients who screen positive for depression and high-risk drinking. BMC Health Services Research, 22, 1123. 10.1186/s12913-022-08408-1 [PubMed: 36064354]

Saitz R (2005). Unhealthy Alcohol Use. New England Journal of Medicine, 352(6), 596–607. 10.1056/NEJMcp042262 [PubMed: 15703424]

Sayre M, Lapham GT, Lee AK, Oliver M, Bobb JF, Caldeiro RM, & Bradley KA (2020). Routine Assessment of Symptoms of Substance Use Disorders in Primary Care: Prevalence and Severity of Reported Symptoms. Journal of General Internal Medicine, 35(4), Article 4. 10.1007/s11606-020-05650-3

Siu AL, US Preventive Services Task Force (USPSTF), Bibbins-Domingo K, Grossman DC, Baumann LC, Davidson KW, Ebell M, Gsarcía FAR, Gillman M, Herzstein J, Kemper AR, Krist AH, Kurth AE, Owens DK, Phillips WR, Phipps MG, & Pignone MP (2016). Screening for Depression in Adults: US Preventive Services Task Force Recommendation Statement. JAMA, 315(4), 380–387. 10.1001/jama.2015.18392 [PubMed: 26813211]

Suen LW, Makam AN, Snyder HR, Repplinger D, Kushel MB, Martin M, & Nguyen OK (2022). National Prevalence of Alcohol and Other Substance Use Disorders Among Emergency Department Visits and Hospitalizations: NHAMCS 2014–2018. Journal of General Internal Medicine, 37(10), 2420–2428. 10.1007/s11606-021-07069-w [PubMed: 34518978]

US Preventive Services Task Force, Curry SJ, Krist AH, Owens DK, Barry MJ, Caughey AB, Davidson KW, Doubeni CA, Epling JW, Kemper AR, Kubik M, Landefeld CS, Mangione CM, Silverstein M, Simon MA, Tseng C-W, & Wong JB (2018). Screening and Behavioral Counseling Interventions to Reduce Unhealthy Alcohol Use in Adolescents and Adults: US Preventive Services Task Force Recommendation Statement. JAMA, 320(18), 1899–1909. 10.1001/jama.2018.16789 [PubMed: 30422199]

Williams EC, Bryson CL, Sun H, Chew RB, Chew LD, Blough DK, Au DH, & Bradley KA (2012). Association between alcohol screening results and hospitalizations for trauma in Veterans Affairs outpatients. The American Journal of Drug and Alcohol Abuse, 38(1), 73–80. 10.3109/00952990.2011.600392 [PubMed: 21797815]

Williams EC, McGinnis KA, Bobb JF, Rubinsky AD, Lapham GT, Skanderson M, Catz SL, Bensley KM, Richards JE, Bryant KJ, Edelman EJ, Satre DD, Marshall BDL, Kraemer KL, Blosnich JR, Crystal S, Gordon AJ, Fiellin DA, Justice AC, & Bradley KA (2018). Changes in alcohol use associated with changes in HIV disease severity over time: A national longitudinal study in the Veterans Aging Cohort. Drug and Alcohol Dependence, 189, 21–29. 10.1016/j.drugalcdep.2018.04.022 [PubMed: 29859388]

Williams EC, McGinnis KA, Rubinsky AD, Matson TE, Bobb JF, Lapham GT, Edelman EJ, Satre DD, Catz SL, Richards JE, Bryant KJ, Marshall BDL, Kraemer KL, Crystal S, Gordon AJ, Skanderson M, Fiellin DA, Justice AC, & Bradley KA (2021). Alcohol Use and Antiretroviral Adherence Among Patients Living with HIV: Is Change in Alcohol Use Associated with Change in Adherence? AIDS and Behavior, 25(1), 203–214. 10.1007/s10461-020-02950-x

Williams EC, McGinnis KA, Tate JP, Matson TE, Rubinsky AD, Bobb JF, Lapham GT, Edelman EJ, Catz SL, Satre DD, Bryant KJ, Marshall BDL, Kraemer KL, Bensley KM, Richards JE, Skanderson M, Justice AC, Fiellin DA, & Bradley KA (2019). HIV Disease Severity is Sensitive to Temporal Changes in Alcohol Use: A National Study of VA Patients with HIV. Journal of Acquired Immune Deficiency Syndromes (1999), 81(4), 448–455. 10.1097/QAI.0000000000002049 [PubMed: 30973541]

Witkiewitz K, Hallgren KA, Kranzler HR, Mann KF, Hasin DS, Falk DE, Litten RZ, O'malley SS, & Anton RF (2017). Clinical validation of reduced alcohol consumption after treatment for alcohol dependence using the World Health Organization risk drinking levels. Alcoholism: Clinical and Experimental Research, 41(1), 179–186. [PubMed: 28019652]

Witkiewitz K, Heather N, Falk DE, Litten RZ, Hasin DS, Kranzler HR, Mann KF, O'Malley SS, & Anton RF (2020). World Health Organization risk drinking level reductions are associated with improved functioning and are sustained among patients with mild, moderate and severe alcohol dependence in clinical trials in the United States and United Kingdom. Addiction, 115(9), 1668–1680. 10.1111/add.15011 [PubMed: 32056311]

Yeung K, Richards J, Goemer E, Lozano P, Lapham G, Williams E, Glass J, Lee A, Achtmeyer C, Caldeiro R, Parrish R, & Bradley K (2020). Costs of using evidence-based implementation strategies for behavioral health integration in a large primary care system. Health Services Research, 55(6), 913–923. 10.1111/1475-6773.13592 [PubMed: 33258127]
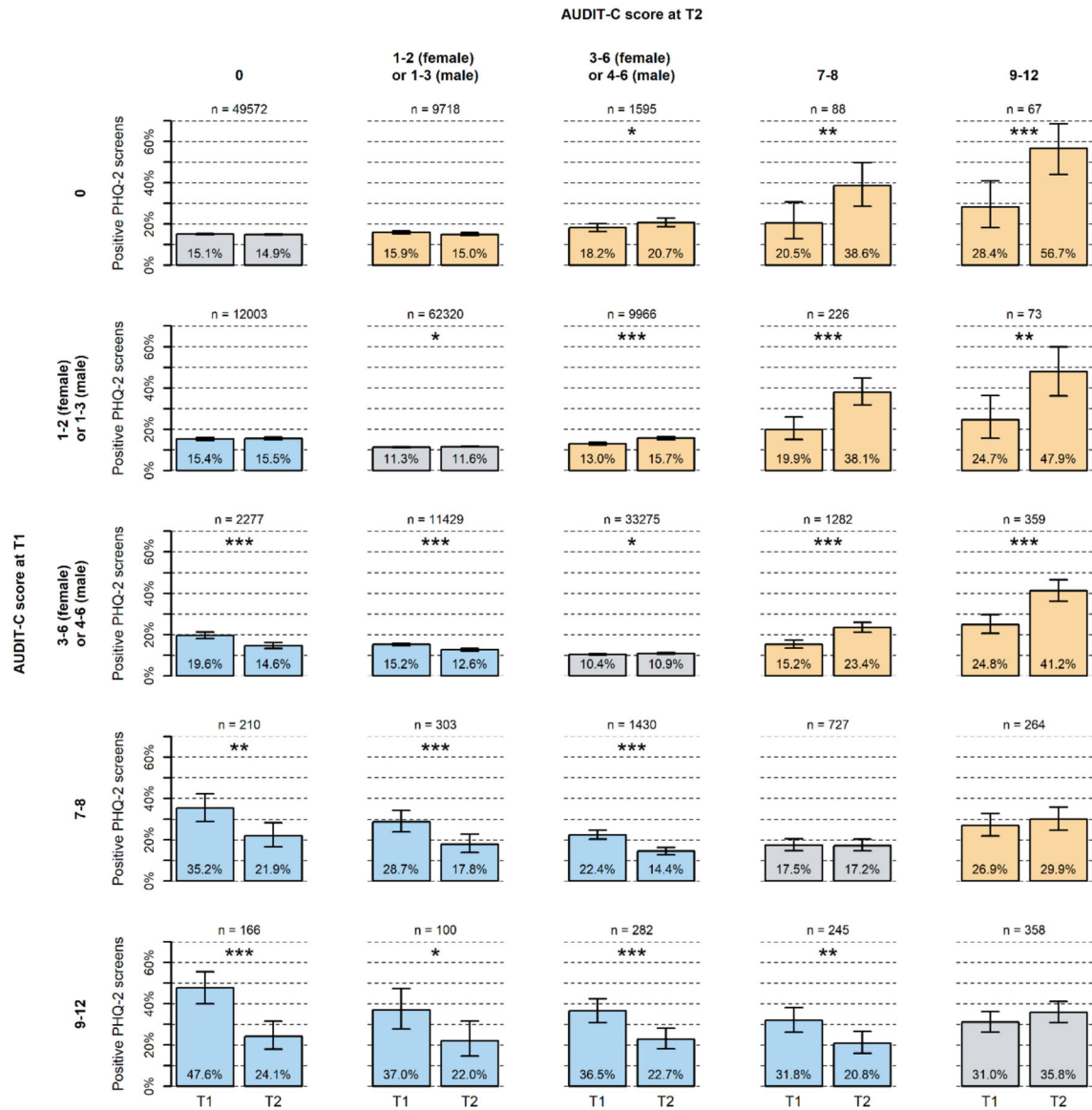
Figure 1. Prevalence of positive PHQ-2 depression screens at T1 and T2, stratified by T1 and T2 AUDIT-C scores.

Gray graphs on the left-to-right diagonal represent subgroups that did not have changes in AUDIT-C score categories from T1 to T2. Orange graphs above the diagonal represent subgroups that had increases in AUDIT-C score categories from T1 to T2. Blue graphs below the diagonal represent subgroups that had decreases in AUDIT-C score categories from T1 to T2. Error bars represent 95% CI's. Asterisks indicate significant changes from T1 to T2 based on McNemar's tests, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

Table 1

Drinking risk categories defined for the current study.

Drinking risk category	AUDIT-C score
No drinking	0
Drinking without unhealthy alcohol use	1–2 (women) or 1–3 (men)
Unhealthy alcohol use, moderate-risk drinking	3–6 (women) or 4–6 (men)
Unhealthy alcohol use, high-risk drinking	7–8
Unhealthy alcohol use, very high-risk drinking	9–12

Note: The drinking risk categories above are derived based on cutoffs for any alcohol use (scores 1), commonly used cutoffs for unhealthy alcohol use that typically warrants brief intervention within primary care (scores 3 for women or 4 for men) (Bradley et al., 2007), high-risk drinking that is typically followed by assessment of AUD symptoms in the KP Washington health system (scores 7) (Sayre et al., 2020), and scores that reflect exponentially increasing levels of alcohol consumption (Rubinsky et al., 2013) associated with markedly increased risk for AUD and adverse health outcomes (scores 9–12) (Au et al., 2007; Bradley et al., 2016; Jack et al., In Press; Rubinsky et al., 2012; Williams et al., 2012).

Table 2

Characteristics of patients meeting study inclusion criteria (N = 198,335).

		%
Age	18–24	6.6
	25–44	26.2
	45–64	36.6
	65+	30.6
Sex	Female	61.9
	Male	38.1
Race	American Indian/Alaska Native	0.6
	Asian	8.9
	Black or African American	4.6
	Hispanic or Latinx	5.3
	Native Hawaiian or Pacific Islander	0.8
	White	72.4
	Another race not listed	1.3
	More than one race	3.0
	Unknown race	3.1
Insurance	Commercial	62.1
	Medicaid	3.4
	Medicare	31.2
Psychiatric diagnoses	Anxiety disorder	19.3
	Attention deficit disorder	2.6
	Bipolar spectrum disorder	2.2
	Depressive disorder	22.5
	Eating disorder	0.5
	Post-traumatic stress disorder	1.9
	Schizophrenia spectrum disorder	0.3
	Other psychosis	0.4
	Other mental health disorder	10.0
	Any mental health disorder above	34.7
	Alcohol use disorder[1]	2.2
	Other substance use disorder[1,2]	2.2
AUDIT-C score at T1	0	30.8
	1–2 (female) or 1–3 (male)	42.6
	3–6 (female) or 4–6 (male)	24.5
	7–8	1.5
	9–12	0.6
AUDIT-C score at T2	0	32.4
	1–2 (female) or 1–3 (male)	42.3
	3–6 (female) or 4–6 (male)	23.5
	7–8	1.3

	%
9–12	0.6

Note.

[1] Excludes disorders that are "in remission".

[2] Excludes tobacco and alcohol use disorders.

Table 3

Within-group risk ratios (RRs) reflecting changes in risk of positive PHQ-2 depression screen from T1 to T2.

		Positive PHQ-2 depression screen	
AUDIT-C score at T1	**AUDIT-C score at T2**	**RR (T2/T1)**	**(95% CI)**
0	0	0.99	(0.96, 1.01)
0	1–2 (female) or 1–3 (male)	0.95	(0.90, 1.00)
0	3–6 (female) or 4–6 (male)	1.14	(1.01, 1.28)
0	7–8	1.89	(1.23, 3.10)
0	9–12	2.00	(1.44, 3.08)
1–2 (female) or 1–3 (male)	0	1.01	(0.96, 1.06)
1–2 (female) or 1–3 (male)	1–2 (female) or 1–3 (male)	1.03	(1.00, 1.05)
1–2 (female) or 1–3 (male)	3–6 (female) or 4–6 (male)	1.21	(1.14, 1.28)
1–2 (female) or 1–3 (male)	7–8	1.91	(1.47, 2.60)
1–2 (female) or 1–3 (male)	9–12	1.94	(1.29, 3.21)
3–6 (female) or 4–6 (male)	0	0.74	(0.66, 0.83)
3–6 (female) or 4–6 (male)	1–2 (female) or 1–3 (male)	0.83	(0.79, 0.88)
3–6 (female) or 4–6 (male)	3–6 (female) or 4–6 (male)	1.05	(1.01, 1.09)
3–6 (female) or 4–6 (male)	7–8	1.54	(1.34, 1.77)
3–6 (female) or 4–6 (male)	9–12	1.66	(1.40, 2.01)
7–8	0	0.62	(0.46, 0.82)
7–8	1–2 (female) or 1–3 (male)	0.62	(0.47, 0.81)
7–8	3–6 (female) or 4–6 (male)	0.64	(0.56, 0.74)
7–8	7–8	0.98	(0.82, 1.18)
7–8	9–12	1.11	(0.89, 1.40)
9–12	0	0.51	(0.37, 0.65)
9–12	1–2 (female) or 1–3 (male)	0.59	(0.38, 0.87)
9–12	3–6 (female) or 4–6 (male)	0.62	(0.49, 0.77)
9–12	7–8	0.65	(0.50, 0.84)
9–12	9–12	1.15	(0.98, 1.36)

Note. PHQ-2 screening scores 3 are considered positive. RR values reflect the relative change in risk from T1 to T2 (i.e., T2 / T1).