Concurrent validity of an Estimator of Weekly Alcohol Consumption (EWAC) based on the Extended AUDIT

Table of Contents

# Abstract

**Background and Aims:** The 3-question Alcohol Use Disorders Identification Test (AUDIT-C) is commonly employed in healthcare to screen for levels of alcohol consumption. AUDIT-C scores (0–12) have no direct interpretation and do not provide information on alcohol intake, an important variable for behaviour change. The study aimed to (a) develop a continuous metric from the Extended AUDIT-C, offering equivalent accuracy, and providing a direct estimator of weekly alcohol consumption (EWAC); (b) evaluate the EWAC’s bias and error using the Graduated-Frequency (GF) questionnaire as a reference standard of alcohol consumption.

**Design:** Diagnostic evaluation using cross-sectional data from a nationally-representative survey.

**Settings:** Community-dwelling households in England.

**Participants:** 22,404 household residents aged 16 years reporting drinking alcohol at least occasionally.

**Measurements:** Computer-assisted personal interviews consisting of (a) Extended AUDIT and (b) GF. The primary outcomes were: mean deviation <1 UK unit (metric of bias); root mean squared deviation <2 UK units (metric of total error) between EWAC and GF. The secondary outcome was an equivalent Receiver Operating Characteristics’ area under the curve for predicting alcohol consumption in excess of 14 and 35 UK units compared to AUDIT-C and AUDIT.

**Findings:** EWAC had a positive bias of 0.2 UK units [95% confidence interval: 0.04, 0.3] compared to GF. Deviations were skewed: while the mean error was ±11 UK units/week [9.8, 12.0], in half of participants the deviation between EWAC and GF was between 0 and ±2.1 UK units/week. EWAC predicted consumption in excess of 14 UK units/week with a significantly greater area under the curve (0.921 [0.917, 0.925]) than AUDIT-C (0.871 [0.866, 0.876]) or the full AUDIT (0.854 [0.849, 0.860]).

**Conclusions:** The EWAC <https://ewac.netlify.app> is designed to estimate weekly alcohol consumption using answers to the Extended AUDIT-C questionnaire. Using the detailed GF as a reference standard, the EWAC met the targeted bias tolerance. Its accuracy was superior to that of both AUDIT-C and the full AUDIT in relation to consumption thresholds, making it a reliable complement to the Extended AUDIT-C for health promotion interventions.

# Keywords

alcohol consumption; self report; alcohol use disorder; screening programs, diagnostic; preventive health services

# Introduction

Alcohol consumption is responsible for 5.3% of all deaths and 5.0% of disability-adjusted life years [[1](#ref-Shield2020)]. This burden extends far beyond those categorised with *alcohol use disorders*, as defined in the International Statistical Classification of Diseases (ICD-10 F10.1/F10.2; ICD-11 QE10/6C40.1) or the Diagnostic and Statistical Manual of Mental Disorders [[2](#ref-DSM5)]. Clinical guidelines aim to prevent [[3](#ref-NICE-PH24)], treat and reduce [[4](#ref-NICE-CG115)] harm from alcohol consumption, and recommend systematic screening for alcohol consumption using a validated diagnostic tool. However, the significant conceptual differences in the diagnostic classifications above highlight the complexity in finding ways of accurately diagnosing, measuring, and communicating harm internationally [[5](#ref-Saunders2019)].

A global standard has emerged in the Alcohol Use Disorders Identification Test (AUDIT) [[6](#ref-Babor2001)], a 10-item questionnaire. Its first three questions focussing on consumption (AUDIT-C) have equivalent predictive capability [[7](#ref-Bush1998)], and are easy to use for patients and clinicians alike, making AUDIT-C an attractive choice for alcohol interventions in healthcare and other settings. Yet, AUDIT-C exhibits two limitations:

1. **Ceiling effect:** the AUDIT-C’s maximum response options for alcohol consumption frequency (‘four or more times per week’) and quantity (‘10 or more drinks’) are heavily right-censored. This provides little discriminant capability among individuals scoring high. Moreover, due to these limited response options, the AUDIT-C is also not responsive to changes in individuals with a high baseline score, even if they succeed in reducing their consumption by as much as thirty percent (eg. frequency of drinking down from 7 to 5 days or quantity down from 16 to 11 drinks per day).
2. **AUDIT score intelligibility:** AUDIT-C produces an ordinal score (range: 0–12). In spite of predicting many dimensions of alcohol risk, including alcohol consumption [[8](#ref-Rubinsky2013)], when presented alone this score does not facilitate intuitive interpretation to patients and healthcare workers alike, as it is not a direct measure of consumption. It must be interpreted by reference to cut-offs informed by diagnostic accuracy studies [[9](#ref-DeMeneses-Gaya2009)], and set with great variation internationally [[10](#ref-Nadkarni2019)]. Although categorical risk categories can be derived from the AUDIT-C (e.g. ‘low’, ‘increasing’ and ‘high’ risk [[11](#ref-Lavoie2010)]), they are also not easy to interpret in relation to epidemiological evidence on dose-response associations between alcohol and a wide range of medical conditions [[12](#ref-Corrao2004)]. The fact that the AUDIT-C score is not easy to interpret poses challenges in the delivery of brief interventions. Healthcare workers lack confidence in discussing alcohol consumption [[13](#ref-Johnson2010)–[15](#ref-McCormick2006)], and the abstract AUDIT-C score requires specific training for healthcare staff on how to feed back results [[13](#ref-Johnson2010)]. This lack of intelligibility also poses challenges in embedding AUDIT-C and communicating its result in self-administered behaviour change interventions such as web apps [[16](#ref-Beyer2018)].

The first of these limitations has been addressed by a recently-introduced variation to the questionnaire: the ‘Extended AUDIT-C’ contains a greater range of response options on quantity and frequency (Table 1). It has been used in UK research as part of two trials [[17](#ref-Kaner2013c),[18](#ref-Crane2018)] and one continuous household survey [[19](#ref-Beard2015a)] to capture greater information on the higher risk drinkers, and detect reductions in consumption that could not have been measured with the right-censored AUDIT-C.

Table 1: Comparison of AUDIT-C and Extended AUDIT-C (questions 1 and 2)

<<INSERT TABLE 1 HERE>>

The present study proposes to address the second of these limitations, by developing and validating a fast and easy-to-complete Estimator of Weekly Alcohol Consumption (EWAC) computed from the Extended AUDIT-C. The EWAC is designed to offer equivalent diagnostic capability in relation to alcohol use disorders (being based on the AUDIT-C), while providing a continuous and more directly interpretable metric through alcohol consumption. This should facilitate the delivery of brief interventions, since measuring alcohol consumption is a crucial part of some of the most common behaviour change techniques (self-monitoring, feedback on behaviour, social comparison) in self-administered [[20](#ref-Kaner2017)] and clinician-administered [[21](#ref-Michie2012)] interventions. The EWAC also facilitates understanding alcohol use disorders and alcohol-related harm across a wide spectrum of levels of consumption which cannot be reduced to a dichotomy [[22](#ref-Vinson1997)]. It has instead been argued [[23](#ref-Rehm2013a),[24](#ref-Nutt2014)] that focusing attention on monitoring and reducing alcohol consumption across the spectrum of use reduces stigma and can facilitate alcohol harm prevention throughout the population.

This structure of the paper is as follows. The paper reports the development of the EWAC, which is the product of quantity (AUDIT-Q1) by frequency (AUDIT-Q2) with adjustment for occasional intense drinking (AUDIT-Q3). Coefficients are estimated in a hierarchical Bayesian item response model using Extended AUDIT-C and Graduated-Frequency (GF) data from the Alcohol Toolkit Study [[19](#ref-Beard2015a)], a large panel survey of community households in England. The same data are then used to test metrics of bias and error, namely the mean deviation and the root mean squared deviation between EWAC and GF. Multivariate regression models of the same metrics are used to detect potential subgroup variations in bias and error in the EWAC. Finally, tests of the equivalence of the full Receiver Operating Characteristics Area Under the Curve of EWAC compared with AUDIT-C and the full AUDIT in predicting alcohol consumption in excess 112g/week (14 UK units) and 280g/week (35 UK units) respectively are reported.

# Methods

## Design

This paper develops and validates an Estimator of usual Weekly Alcohol Consumption in UK units (EWAC) based on the Extended AUDIT-C. Neither the AUDIT-C nor the Extended AUDIT-C provide a direct measure of alcohol consumption. However, the product of frequency of drinking (AUDIT-Q1) and usual quantity of drinking per occasion (AUDIT-Q2) can be used to estimate usual average weekly alcohol consumption, with adjustment for occasional heavy use (AUDIT-Q3), following methods developed for quantity-frequency-variability instruments [[25](#ref-Lemmens1992)].

In practice, for every individual , the EWAC is computed as the product of and (AUDIT questions 1 and 2 respectively) adjusted with the frequency of intense drinking (AUDIT-Q3):

where denotes the average units of alcohol consumed in an intense drinking day. Coefficients and are unknown. In this study, two sets of potential coefficients are evaluated using:

* the AUDIT response item interval midpoint: for example, 2.5 for ‘2 to 3 times per week’ in AUDIT question 2
* coefficients estimated empirically from a sample of individuals with measurements of Extended AUDIT-C and GF, using a hierarchical Bayesian response model with the estimating equation , where denotes a vector of independently normally distributed errors. We set parabola-shaped informative priors on coefficients in the form of distributions bounded by the response item interval and centred on its midpoint. A normal prior was used for . Priors were subsequently tuned to improve convergence. Details on model fitting, convergence evaluation and prior tuning is reported in supplementary information S1.

## Participants

Data were obtained from waves 110–133 of the Alcohol Toolkit Study (ATS), a computer-assisted personal interview of residents of private English households aged 16 years and over [[19](#ref-Beard2015a)]. Data collection took place between November 2015 and October 2017, when GF questions were included. Participants were included in the analysis if they completed both the Extended AUDIT and the Graduated Frequency (GF) alcohol consumption questionnaire. Out of a total 40,832 participants, 14,408 (35%) reported ‘never’ consuming alcohol in the AUDIT-Q1 and were not asked any further AUDIT or GF questions. They were thus excluded. Although the ATS included a shorter follow-up six months after the initial interview, corresponding data were not used as they did not include GF measurements.

## Measures

Index measures underpinning the EWAC were the response items to the three questions making up the Extended AUDIT-C (supplementary information S2), in which participants are prompted to describe their drinking *during the last 6 months*.

The reference standard used is the ATS’ Graduated-Frequency (GF) schedule (supplementary information S3), in which participants described how many times they consumed given quantities of alcohol *during the last 4 weeks* [[26](#ref-Greenfield2000)]. A longstanding obstacle in alcohol research and treatment is that there is no undisputed ‘gold standard’ or biomarker for objectively determining recent levels of alcohol consumption. Instead, a number of instruments exist which measure self-reported alcohol consumption with varying validity and reliability over different time periods. Compared with population-level sales and production data, survey data tends to under-report consumption levels [[27](#ref-Livingston2015)]. Comprehensive reviews [[28](#ref-Heeb2005)–[32](#ref-Stockwell2016)] indicate that yesterday recall and prospective diaries tend to record higher (and more accurate) alcohol consumption by minimising recall bias, followed by GF measures. Lower levels are recorded with quantity-frequency measures. The GF schedule has the advantage of measuring occasional heavy consumption, which can constitute an important proportion of all alcohol consumed. Despite being commonly used in population surveys, GF measures are not widely employed in clinical practice.

Other reference data were used to cross-examine the EWAC’s coverage of retail sales data and population-wide cumulative distribution in England. We used 2011 and 2014 data as these were the years for which all required measures were available. Per-capita alcohol retail sales for 2014 [[33](#ref-PHE2017)] were used, which measured the ratio of all alcohol produced or processed in the UK, as well as alcohol imported into the UK for sale and consumption, to the mid-year population estimate of the resident population aged 18 years and over. Alternative estimates of the alcohol consumption mean, median and cumulative distribution in England were obtained from 8,610 residents of private English households aged 16 years and over participating in the 2011 Health Survey for England (HSE) [[34](#ref-NatCenSocialResearch2013)]. On that particular year, the HSE included (a) a computer-assisted interviewer led beverage-specific quantity-frequency questionnaire; (b) a yesterday recall question [[32](#ref-Stockwell2016)]; and (c) a prospective 7-day diary [[35](#ref-Boniface2014)].

## Missing data

Out of 26,424 participants who did not answer ‘never’ to AUDIT-Q1, 4,020 (15.2%) did not have a valid GF alcohol consumption record and were excluded. In total, 22,404 valid observations remained for the diagnostic analysis, in which missing GF data were assumed to be missing at random conditionally on the Extended AUDIT-C responses. In the subgroup analysis, a further 530 respondents (2.4%) were assumed to have data missing at random and were excluded.

## Analyses

A pre-registered protocol for this study is available online [[36](#ref-Dutey2018)]. All results are reported in UK alcohol units (8g or 10mL of pure alcohol). Analyses are conducted in R [[37](#ref-RCoreTeam2017)] using packages tidyverse and rstan [[38](#ref-package-tidyverse),[39](#ref-package-rstan)]. Computer scripts for all analyses are available online [[40](#ref-Dutey2020)].

### Overall bias and error

The agreement between the EWAC and the GF was quantified with the help of two deviation metrics:

* bias is estimated using the **mean deviation** to the reference standard $\text{MD} = n^{-1} \sum\_{i=1}^{n}{(\rm{EWAC}\_i - \rm{GF}\_i )}$. We test the hypothesis that the MD is greater than 1 UK unit using a two-sided -test.
* precision is estimated using the **root mean squared deviation** $\text{RMSD} = \sqrt{n^{-1} \sum\_{i=1}^{n}{( \rm{EWAC}\_i - \rm{GF}\_i )^2}}$, a measure of total error: it captures both bias and random deviation from the reference standard. For example, an RMSD of 2 signifies that the EWAC is on average with 2 UK units of the reference standard. We test the hypothesis that the RMSD is greater than 2 UK units using a one-sided homogeneity test.

Two sets of coefficients were compared: AUDIT response item interval midpoint; and coefficients estimated using a statistical model (see supplementary information S1). The best set of coefficients was retained for the remainder of the analysis.

### Subgroup bias and error

Multivariate regression models tested whether the EWAC’s validity varies across population subgroups:

* the simple deviation was regressed in a linear model to test for subgroup differences in MD
* the squared deviation $(\rm{EWAC}-\rm{GF})^2$ was regressed in a log-transformed linear model to test subgroup differences in the geometric mean squared deviation. Model coefficients were then back-transformed (square root of the exponential) into relative RMSD estimates interpreted as the ratio of the subgroup RMSD to the reference category RMSD, a ratio >1 indicating worse precision than in the reference category.

Both models included the following predictors: sex by age group; ethnic group; highest educational qualification; religion; smoking status. Additional models were fitted solely in respondents with an AUDIT-C score of 5 or more or an AUDIT score of 8 or more, for whom additional characteristics were recorded during interview: favourite drink (beer; wine; spirits alone; mixed spirits; cider; other); and whether the respondent had attempted to restrict alcohol intake in the last 12 months (e.g. by drinking less, choosing lower strength alcohol or using smaller glasses).

### Receiver Operating Characteristics

We tested the EWAC’s superiority to the traditional AUDIT and AUDIT-C scores in predicting consumption in excess of 14 or 35 UK units per week. These correspond to UK thresholds for characterising alcohol use as ‘increasing risk’ (predicted by an AUDIT-C score of 5-7), and ‘higher risk’ (AUDIT-C score 8) which is above 35 units for women and 50 units for men [[11](#ref-Lavoie2010)]. We tested the hypothesis that the EWAC’s receiver operating characteristics’ full area under the curve (AUC) is greater than that of the AUDIT-C and the full AUDIT scores using nonparametric paired AUC tests [[41](#ref-Delong1988)]. AUDIT-C and AUDIT scores were calculated from the Extended AUDIT by capping the contribution of each response item to 4.

### Aggregate concurrent validity

We compared the EWAC with other population-level alcohol consumption estimates in England in residents aged 18 years and over. We computed the empirical cumulative distributions of alcohol consumption given by (1) the EWAC computed in the ATS; (2) the GF estimator in the ATS; (3) the beverage-specific estimator in HSE in 2011; (4) the prospective diary estimator in HSE 2011. Survey weights were used: in analyses (1-3), poststratification weights estimated using calibration and age-sex mid-year population estimates; in analysis (4), similar poststratification weights adjusted for self-selection into participation to the prospective diary data collection. We report the percentage of total alcohol sales for England accounted for by each method, using both on-trade and off-trade 2014 sales estimates from Public Health England [[33](#ref-PHE2017)].

# Results

## Overall bias and accuracy

The first step involved choosing a set of coefficients to compute the EWAC.

Using interval midpoint coefficients, the EWAC Pearson’s correlation with GF was = 0.69 [95% CI: 0.69, 0.70]. The mean deviation (MD) was 0.7 UK alcohol units/week [95% CI: 0.5, 0.9]. This bias is smaller than the preregistered 1-unit bias tolerance ( = 1.000). The root mean squared deviation (RMSD) estimate of 12.3 units [95% CI: 11.2, 13.2] was significantly greater than the pre-registered 2-unit total error tolerance ( < 0.001).

Coefficients estimated empirically (supplementary information S4) provided small improvements: the EWAC’s Pearson’s correlation with GF was estimated at = 0.71 [0.71, 0.72] (Kendall’s rank correlation = 0.63) and MD = 0.2 [95% CI: 0.04, 0.3] ( = 1.000). With RMSD = 10.9 [95% CI: 9.8, 12.0], precision remained statistically significantly greater than 2 ( < 0.001), suggesting that the EWAC falls on average 11 units away from the GF reference standard.

The RMSD masked a dispersed and skewed distribution of error. Table 2 shows that, for 50% of participants, the EWAC fell within 2 UK units of the GF weekly consumption estimate. Thus, an interval estimate defined as the EWAC 2 units (e.g. ‘10 to 14 units’) would contain the reference standard for half of individuals, while an interval estimate defined as the EWAC 3 units (e.g. ‘9 to 15 units’) would contain the reference standard for 60% of individuals.

Table 2: Percentiles of the absolute deviation between EWAC and GF schedule (n = 22,404)

10%

20%

30%

40%

50%

60%

70%

80%

90%

95%

99%

0.4

0.7

1.0

1.5

2.1

3.0

4.2

6.2

10.6

17.0

38.7

Figure 1 compares individual EWAC and GF values. Lines of best fit depart from the diagonal, demonstrating the EWAC’s small positive bias (MD > 0) is not consistent. The plots indicate a slight positive bias for consumptions up to 10-14 units/week, then a slight negative bias above this threshold. The EWAC loses some of its granularity above 70 units/week (99th percentile of its distribution), where it provides just 6 possible values (82; 83; 92; 93; 100; 125 units/week; see Figure 1(b)), although this remains superior to what is achievable with the AUDIT-C.

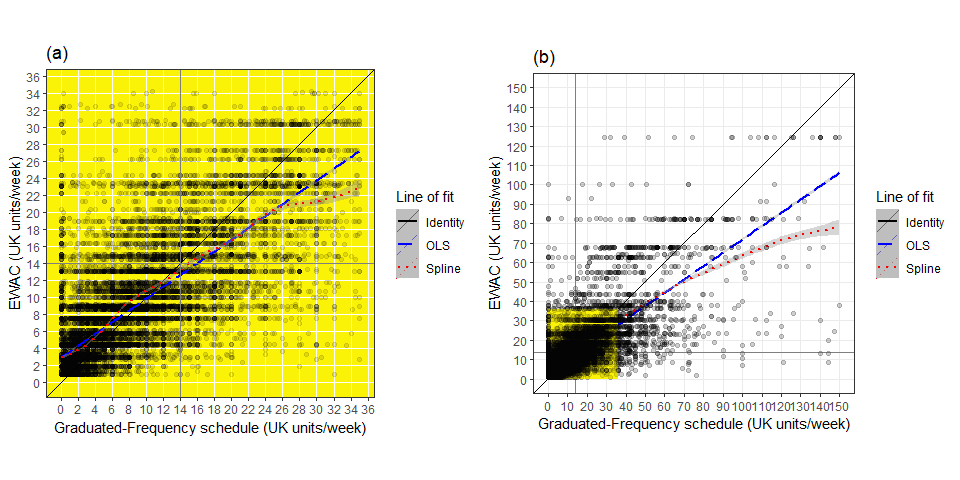


Figure 1: Plots of EWAC against GF in (a) low/increasing risk ATS respondents (n=21,338) and (b) all ATS respondents (n=22,373)

## Subgroup accuracy

Next, the MD and RMSD were regressed against respondent characteristics in order to identify potential subgroup differences in bias or precision (supplementary information S5, Table A). The reference category was: females aged between 25 and 34 years of White ethnicity without educational qualifications, who never smoked. Although the model predictors explained a very modest proportion of both MD and RMSD ( statistics < 2%), several subgroups had different MD and RMSD.

Figure 2 summarises the MD for subgroups whose predicted MD was either above 1 or below -1; and whose coefficients had a -value below 0.05. Respondents of Black, Other, and White Other ethnic groups had significantly overestimated EWACs: their MDs were respectively 4.8 units [95% CI: 2.1, 7.5]; 5.9 units [1.6, 10.1] and 1.6 [0.2, 3.0] in excess of the reference MD. The MDs of respondents aged 55 to 64 years, or 75 years and over respectively had MDs 2.2 units [0.5; 3.9] and 4.2 units [0.9; 7.6] in excess of the reference MD. We found no significant evidence of an effect of favourite drink or attempts to reduce alcohol intake in the past year in increasing-risk drinkers (supplementary information S5, Table B)

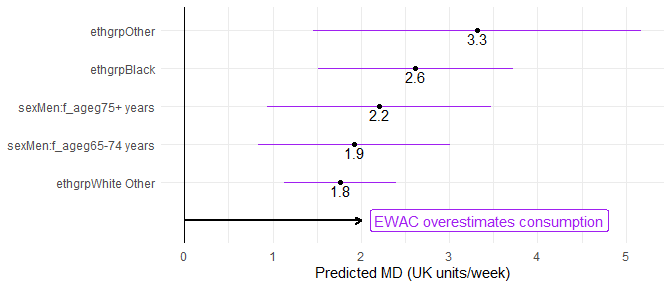


Figure 2: Forest plot of modelled MD for selected subgroups

Figure 3 shows subgroups with an RMSD significantly different from the rest of the population (); and estimated to be 20% greater or smaller than the RMSD of the reference category. This shows that RMSD was 58% [95% CI: 50; 67%] greater in current smokers, 34% [14; 56%] greater in respondents who stopped smoking in the past year, and 23% greater [17; 30%] in respondents who stopped smoking over a year ago. It was also 44% greater [29; 60%] greater in men and 34% greater [19; 50%] in respondents aged 16 to 24 years. In contrast, error was 35 to 70% smaller in White Other, Black and Other ethnic groups.

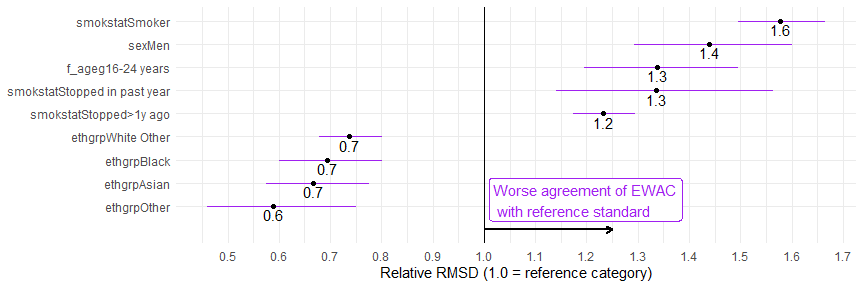


Figure 3: Forest plot of RMSD ratio (selected subgroups to reference category)

Figure 4 visualises the same analysis in increasing-/higher-risk drinkers (AUDIT-C 5 or AUDIT 8). With a higher mean alcohol consumption, results differed from the overall picture presented in Figure 3. The RMSD of respondents favouring mixed spirits had an RMSD 23% [95% CI: 9.6; 38] smaller than the reference category. Educational qualifications significantly improved the agreement between EWAC and the reference standard. School and degree-level qualifications reduced RMSD by 24% [12; 37%] and 37% [23; 52%] respectively, suggesting that respondents may have better recall and clarity over alcohol beverage content. Conversely, the RMSD of respondents who attempted to reduce their alcohol consumption was 23% [16; 30%] larger than the reference category. This is attributable in part to the greater dispersion of values in groups with higher alcohol consumption.

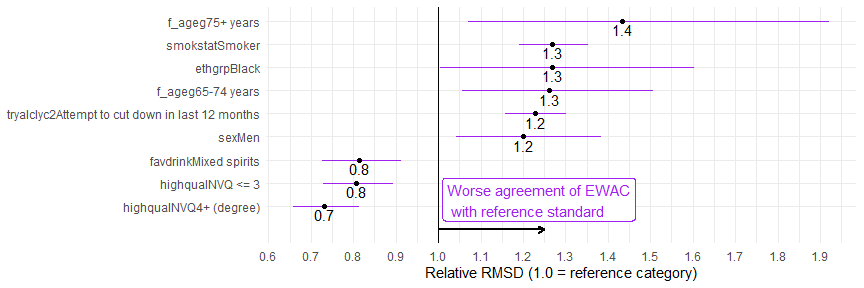


Figure 4: Forest plot of RMSD ratio (selected subgroups to reference category) in respondents with a hazardous/harmful alcohol use (AUDIT-C>=5 or AUDIT>=8

## Receiver Operating Characteristics

We examined the EWAC’s ability to predict consumption exceeding 14 or 35 UK units/week. Receiver operating characteristics curves (available in supplementary information S6) are summarised with respect to the full area under the curve as well as sensitivity/specificity for the best thresholds in Table 3 and Table 4.

**Lower-risk drinking level (14 UK units/week)**: EWAC increased the full area under the curve by 5 percentage points compared with the AUDIT-C score ( < 0.001) and by 7 percentage points compared with the full AUDIT score ( < 0.001), The cut-off maximising the sum of specificity and sensitivity on the EWAC was 10 units/week. The sensitivity at this threshold was identical to AUDIT-C, but specificity was increased by 13 percentage points. Using the nominal cut-off of 14 units/week on the EWAC increased specificity to 0.922, at the cost of a smaller sensitivity of 0.705.

Table 3: Receiver operating characteristics of AUDIT-C score and EWAC for consumption >= 14 UK units or 112g/week (n = 22,404)

Index test

AUC

95% CI

Best threshold

Sensitivity

Specificity

AUDIT-C score

0.871

[0.866, 0.876]

4.5

0.882

0.684

Full AUDIT score

0.854

[0.849, 0.86]

5.5

0.791

0.751

EWAC

0.921

[0.917, 0.925]

10.0

0.876

0.816

*Note*: The best threshold refers the cut-off value that maximises the sum of sensitivity and specificity.

**Higher-risk drinking level (35 UK units/week)**: EWAC provides a small increase in the full area under compared with the AUDIT-C score ( < 0.001) and the full AUDIT score ( < 0.001). The best cut-off for detecting consumption in excess of 35 units/week using the EWAC was 17.

Table 4: Receiver operating characteristics of AUDIT-C score and EWAC for consumption >= 35 UK units or 280g/week (n = 22,404)

Index test

AUC

95% CI

Best threshold

Sensitivity

Specificity

AUDIT-C score

0.913

[0.904, 0.921]

6.5

0.862

0.809

Full AUDIT score

0.901

[0.893, 0.909]

6.5

0.903

0.741

EWAC

0.936

[0.929, 0.943]

16.8

0.865

0.866

*Note*: The best threshold refers the cut-off value that maximises the sum of sensitivity and specificity.

## Empirical distribution functions

Table 5 provides estimates of adult residents’ total alcohol consumption in England using four different estimators, and compares them with alcohol retail sales. The HSE schedules provide the highest estimates of alcohol consumption and coverage of sale statistics. The EWAC estimates of total consumption represent 71% of the HSE prospective diary, and 48% of retail sales.

Table 5: Summary statistics on alcohol consumption in England in residents aged 18 years and over (excluding abstainers)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Study | Mean (units/week) | Median | Variance | N | % of alcohol sold |
| HSE beverage-specific QF | 14.0 | 7.3 | 474.6 | 6,545 | 72.6 |
| HSE prospective diary | 13.0 | 8.0 | 264.7 | 4,640 | 67.6 |
| ATS GF | 8.5 | 5.1 | 242.0 | 22,136 | 44.0 |
| ATS EWAC | 9.3 | 5.2 | 148.9 | 25,882 | 48.0 |
| Retail sales | 19.3 | – | – | – | – |

Figure 5 suggests that the EWAC, like the ATS GF, estimates a higher prevalence of low-risk ( 14 units/week) and increased-risk alcohol use than HSE. In contrast, HSE measures a higher prevalence of alcohol consumption in excess of 50 units/week. This difference may be due to a combination of factors, including differences in sampling coverage, non-response bias, or measurement error in the alcohol schedules.

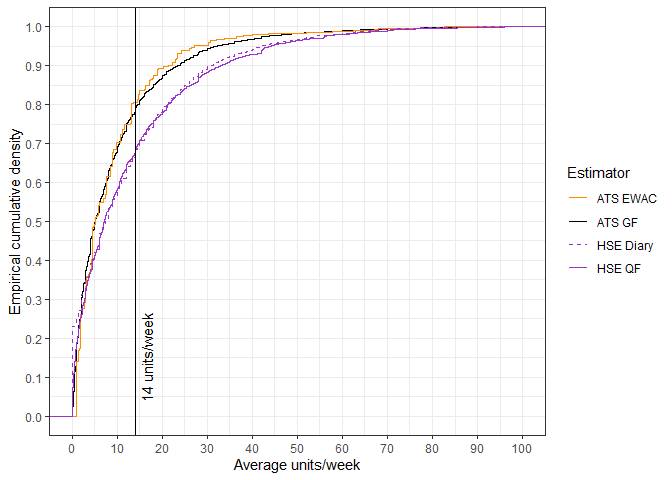


Figure 5: Empirical cumulative distribution function of weekly alcohol consumption in England according to four alcohol schedules in residents aged 18 years and over

# Discussion

## Main findings

This paper reports the development of a continuous metric of estimated weekly alcohol consumption computed from the 3-question Extended AUDIT-C. In a national survey of alcohol consumption in community dwelling residents in England, the resulting metric estimated usual alcohol consumption with a mean precision of 11 units/week when compared to the GF alcohol schedule. Bias was mostly consistent across subgroups examined (age/sex, education, smoking status, religion), with one exception. We found strong evidence that EWAC overestimates alcohol consumption by 2-3 units/week in Black and Other ethnic groups. Despite this, total overall error was smaller in Black, Asian and Other ethnic groups. Variation in diagnostic accuracy across ethnic groups has been previously been noted in a US investigation of AUDIT-C [[42](#ref-Frank2008)].

## Strengths and limitations

To the best of our knowledge, this paper is the first to (a) develop an EWAC using a well-accepted and validated alcohol screening tool such as the AUDIT; and (b) quantify its bias and precision with respect to a continuous measure of alcohol consumption. One US study [[8](#ref-Rubinsky2013)] previously reported mean consumption by AUDIT-C score, yet without quantifying bias or precision of such a measure. Other studies have evaluated the AUDIT-C’s potential in estimating alcohol consumption, but only in relation to predicting consumption in excess of a predefined threshold [[9](#ref-DeMeneses-Gaya2009)]. Such studies achieved AUCs ranging 0.83-0.96. In the present study, we showed the EWAC’s superiority over both the AUDIT-C and the full AUDIT in predicting GF in excess of 14 and 35 units/week. At the 14 unit threshold, the specificity gain from 0.684 (AUDIT-C) to 0.876 places the EWAC with the other best-performing diagnostic tools. Our study provides strong confidence in the internal and external validity of findings in England on account of the large sample size of the ATS, as well as the extensive range of subgroup analyses reported.

Repurposing a well-known tool such as the AUDIT-C has several advantages. It is already translated in many languages and adapted to the varying standard drink sizes adopted internationally [[6](#ref-Babor2001)]. The Extended AUDIT scores can be converted into traditional AUDIT scores by capping items to 4, thereby offering a point of comparison with existing evidence. The AUDIT’s properties are also well understood in diverse contexts and modes of administration as it has been the subject of intense research internationally over the last 30 years. For instance, a previous study found the AUDIT-C to be responsive to changes as small as 70g/week [[43](#ref-Bradley1998)]. We can infer from this past study that the EWAC’s responsiveness to change should be equivalent, if not greater, given the Extended AUDIT-C’s wider range response items than the AUDIT-C.

There are two further strengths of the EWAC, which also highlight the limitations of all screening tools for alcohol consumption and ways of categorising alcohol use.

First the challenges that remain over the conceptualisation of alcohol use disorder itself, which is related to, but not dependent upon the amount of alcohol consumed. Since Jellinek described ‘the disease concept of alcoholism’ sixty years ago [[44](#ref-Jellinek1960)] there have been numerous attempts to categorise the broad range of phenotypes that characterise alcohol use disorders in a way that is valid, reliable, and useful, in the absence of any bio-marker to ‘verify’ the presence of a particular pathology. The EWAC, by limiting itself to an estimation of alcohol consumption is transparent across a wide range of levels of alcohol consumption but does not measure any other factor of this complex and heterogeneous condition [[5](#ref-Saunders2019),[45](#ref-Leggio2009)].

Second, is that the current indirect risk measures (e.g. AUDIT) aid clarity of public health and health service messaging by giving simple cut off points for different levels of alcohol use disorder. However, within these categories substantial variation is hidden, dependent on the characteristics of the population tested (including age, sex, ethnicity and prevalence of AUD), and the type of risk being screened for (e.g. chronic physical versus acute mental harm). (ref Menzies SR, wood et 2018), which may undermine efforts to develop our understanding of relative risks and potential benefits of targeting interventions.

We note other limitations. Findings reported in this paper may not apply to other countries, or to small subpopulations with an atypical alcohol consumption, such as patients seeking care for severe alcohol dependence or alcohol-related liver disease. Like the HSE, ATS does not cover populations excluded from most sampling frames, such as residents of communal and carceral institutions, homeless people, or migrant populations. Independent of coverage, evidence suggests that survey non-respondents may have higher levels of alcohol consumption and harm [[46](#ref-Gorman2014)–[48](#ref-Boniface2017)]. Also, the proportion of respondents classified as non-drinkers (based on AUDIT-Q1) is 10% higher in ATS than in HSE data, where respondents are asked to confirm whether they never drink alcohol, or only drink ‘very occasionally’ [[49](#ref-DeVocht2016)]. All these factors are likely to affect MD and RMSD estimates, and explain some of the discrepancy found in Table 5.

We also note that the selected reference standard has a strong influence on estimates of precision (RMSD). By definition, the reference standard’s own error will inflate the RMSD. This means some of the RMSD is likely to be attributable to error in the reference standard rather than the EWAC. Previous research offers reassurance that the EWAC’s error is comparable to that of other self-reported alcohol questionnaires. A previous US telephone interview study [[50](#ref-Greenfield2009)] measured a Pearson’s correlation coefficient of = 0.86 and 0.87 between drinking frequencies and volumes as measured by a 12-month GF interview and a subsequent 28 day prospective diary. Similarly, Hilton [[51](#ref-Hilton1989)] measured a correlation = 0.89 between a 30-day GF and a prospective diary (identical reference periods). Heeb *et al.* [[28](#ref-Heeb2005)] measured a Kendall’s rank correlation = 0.41 between a 7-day GF and a 7-day prospective diary. For reference, our study measured a correlation coefficient = 0.71 and = 0.63 between a 6-month EWAC and a 4-week GF, that is, a similar level of agreement as GF with diaries.

## Potential applications

The EWAC retains the AUDIT’s strengths (speed, accuracy, international standardisation), and is therefore suitable for use in clinical practice. It can be used in brief interventions to feed back a reliable 4-unit wide interval estimate of alcohol consumption (eg: ‘9-13 units/week’ or ‘70-100 g/week’). One such calculator is available online:

<https://ewac.netlify.app>

EWAC is already compatible with medical records information models developed in the Systematised Nomenclature of Medicine Clinical Terms (SNOMED CT, Alcohol intake (observable entity) [[52](#ref-SNOMED-UK)]) and by the English Royal College of Physicians [[53](#ref-RCP2018)]. Such information can have secondary uses as a variable in other disease risk scores. Provided consistent recording over time, this could facilitate the estimation of lifetime exposure to alcohol, an important risk factor for a wide range of medical conditions.

In comparison to an AUDIT-C score, alcohol consumption is an understandable scale with health education and promotion qualities. It has the potential to facilitate uptake with brief interventions targeting skills in recognising the alcohol content of different drinks and drink sizes to reduce consumption. This may be particularly relevant to primary and secondary prevention efforts to provide earlier interventions (identification and brief advice) to increasing-risk drinkers [[11](#ref-Lavoie2010)]. Knowledge of alcohol beverage content is generally poor [[54](#ref-Mongan2015)], and many countries have not adopted a measure of standard drinks [[55](#ref-Kalinowski2016)]. In England, two thirds of drinkers could assess the standard drink equivalent in wine or beer of one alcohol unit [[56](#ref-ONS2009)].

Assessment of alcohol consumption is not well embedded in clinical practice [[57](#ref-Browne2016)]. Reasons include: difficulty in the self-assessment of alcohol consumption, lack of a single biological marker, time pressures, and stigma [[13](#ref-Johnson2010)]. The EWAC calculator fills a gap in the range of resources to empower individuals to understand, monitor and control their alcohol consumption with–or without–the involvement of healthcare professionals.

Nutt, Rehm *et al.* [[24](#ref-Nutt2014),[58](#ref-Rehm2016)] argue that alcohol-related harm is best prevented if individuals know their consumption level, and health professionals in all settings can engage patients effectively to manage risks with evidence-based interventions, in a similar way to other risk factors for disease, for example blood pressure or cholesterol.

The proposed EWAC can fulfil the same alcohol consumption screening functions as the AUDIT-C, while at the same time providing transparent and personalised information to patients in a less stigmatising way than current terminology (e.g. hazardous and harmful drinking): an estimate of current alcohol consumption is more accessible and understandable than an AUDIT risk score.

This dimensional rather than categorical approach to communicating an estimate of alcohol consumption may help to motivate increasing-risk drinkers to monitor and reduce their consumption, while at the same time encouraging low-risk drinkers to maintain this level. The EWAC appears particularly suitable to integrate personalised feedback and monitoring into digital interventions.

# Declarations

## Acknowledgements

This research was supported by the Medical Research Council [grant reference MR/P016960/1]. The ATS data collection was funded primarily by the National Institute for Health Research (NIHR) School for Public Health Research [grant reference SPHR‐SWP‐ALC‐WP5]. and Public Health Research Programme [grant reference 15/63/01]. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

## Ethics

This study was approved by the University of Southampton’s Faculty of Medicine Ethics Committee (ERGO 44682).

## Consent for publication

Not applicable.

## Competing interests

JB has received unrestricted research funding to study smoking cessation from companies who manufacture smoking cessation medications.

# Supplementary information

## Supplementary information S1: Bayesian model report (PDF file)

## Supplementary information S2: ATS Extended AUDIT questionnaire schedule

## Supplementary information S3: ATS Graduated-Frequency questionnaire schedule

## Supplementary information S4: EWAC coefficients (CSV file)

## Supplementary information S5: Subgroup analyses

Table 6: Coefficients of linear regression of the bias and error of EWAC compared with GF in all respondents (n = 21,874)

(EWAC\_QFV - GFMEANWEEKLY)

LOG((EWAC\_QFV - GFMEANWEEKLY)2)

1

2

Constant

0.7 (-0.1, 1.5)

0.7 (0.5, 0.9)\*\*\*

sexMen

-1.2 (-2.0, -0.4)\*\*

0.7 (0.5, 0.9)\*\*\*

f\_ageg16-24 years

0.7 (-0.1, 1.6)

0.6 (0.4, 0.8)\*\*\*

f\_ageg35-44 years

-0.6 (-1.4, 0.2)

0.03 (-0.2, 0.2)

f\_ageg45-54 years

0.03 (-0.7, 0.8)

0.1 (-0.1, 0.3)

f\_ageg55-64 years

0.5 (-0.3, 1.3)

0.1 (-0.1, 0.3)

f\_ageg65-74 years

0.1 (-0.8, 0.9)

-0.2 (-0.4, 0.03)

f\_ageg75+ years

-0.2 (-1.2, 0.8)

-0.3 (-0.5, -0.00)\*

ethgrpWhite Other

1.1 (0.5, 1.7)\*\*\*

-0.6 (-0.8, -0.4)\*\*\*

ethgrpMixed

-0.5 (-1.8, 0.8)

-0.3 (-0.6, 0.1)

ethgrpAsian

0.9 (-0.2, 2.1)

-0.8 (-1.1, -0.5)\*\*\*

ethgrpBlack

2.0 (0.9, 3.1)\*\*\*

-0.7 (-1.0, -0.4)\*\*\*

ethgrpOther

2.7 (0.8, 4.5)\*\*

-1.1 (-1.6, -0.6)\*\*\*

religionChristian

0.00 (-0.3, 0.3)

-0.2 (-0.3, -0.2)\*\*\*

religionMuslim

1.0 (-1.8, 3.7)

-0.5 (-1.2, 0.2)

religionAny other religion

-0.5 (-1.3, 0.3)

-0.03 (-0.2, 0.2)

highqualNVQ < = 3

-0.01 (-0.6, 0.5)

0.01 (-0.1, 0.1)

highqualNVQ4+ (degree)

-0.1 (-0.7, 0.4)

0.1 (-0.02, 0.3)

highqualOther

-0.3 (-1.0, 0.4)

0.01 (-0.2, 0.2)

smokstatStopped> 1y ago

-0.1 (-0.4, 0.3)

0.4 (0.3, 0.5)\*\*\*

smokstatStopped in past year

-0.3 (-1.5, 0.9)

0.6 (0.3, 0.9)\*\*\*

smokstatSmoker

-0.7 (-1.1, -0.2)\*\*

0.9 (0.8, 1.0)\*\*\*

sexMen:f\_ageg16-24 years

-0.1 (-1.3, 1.0)

-0.1 (-0.4, 0.2)

sexMen:f\_ageg35-44 years

0.04 (-1.1, 1.2)

0.00 (-0.3, 0.3)

sexMen:f\_ageg45-54 years

-0.5 (-1.5, 0.6)

0.04 (-0.2, 0.3)

sexMen:f\_ageg55-64 years

-0.1 (-1.2, 1.0)

0.2 (-0.1, 0.5)

sexMen:f\_ageg65-74 years

1.3 (0.2, 2.4)\*

0.3 (0.00, 0.6)\*

sexMen:f\_ageg75+ years

1.5 (0.3, 2.8)\*

-0.2 (-0.5, 0.1)

Observations

21,874

21,874

R2

0.01

0.05

Adjusted R2

0.01

0.05

Residual Std. Error (df = 21846)

11.0

2.9

F Statistic (df = 27; 21846)

5.1\*\*\*

40.0\*\*\*

Notes:

\*P < .05

\*\*P < .01

\*\*\*P < .001

Table 7: Coefficients of linear regression of the bias and error of EWAC compared with GF in respondents with a hazardous/harmful alcohol use (AUDIT-C>=5 or AUDIT>=8; (n = 9,850)

(EWAC\_QFV - GFMEANWEEKLY)

1

2

Constant

0.4 (-1.4, 2.2)

0.4 (-1.4, 2.2)

sexMen

-1.3 (-2.9, 0.3)

-1.3 (-2.9, 0.3)

f\_ageg16-24 years

1.0 (-0.6, 2.6)

1.0 (-0.6, 2.6)

f\_ageg35-44 years

-0.8 (-2.5, 0.9)

-0.8 (-2.5, 0.9)

f\_ageg45-54 years

0.5 (-1.1, 2.2)

0.5 (-1.1, 2.2)

f\_ageg55-64 years

2.2 (0.5, 3.9)\*

2.2 (0.5, 3.9)\*

f\_ageg65-74 years

1.4 (-0.7, 3.4)

1.4 (-0.7, 3.4)

f\_ageg75+ years

4.2 (0.9, 7.6)\*

4.2 (0.9, 7.6)\*

ethgrpWhite Other

1.6 (0.2, 3.0)\*

1.6 (0.2, 3.0)\*

ethgrpMixed

-1.5 (-4.0, 1.0)

-1.5 (-4.0, 1.0)

ethgrpAsian

2.4 (-0.5, 5.2)

2.4 (-0.5, 5.2)

ethgrpBlack

4.8 (2.1, 7.5)\*\*\*

4.8 (2.1, 7.5)\*\*\*

ethgrpOther

5.9 (1.6, 10.1)\*\*

5.9 (1.6, 10.1)\*\*

religionChristian

0.2 (-0.4, 0.8)

0.2 (-0.4, 0.8)

religionMuslim

4.6 (-4.7, 13.9)

4.6 (-4.7, 13.9)

religionAny other religion

-1.7 (-3.4, -0.01)\*

-1.7 (-3.4, -0.01)\*

highqualNVQ < = 3

0.3 (-0.9, 1.4)

0.3 (-0.9, 1.4)

highqualNVQ4+ (degree)

0.1 (-1.2, 1.3)

0.1 (-1.2, 1.3)

highqualOther

-0.8 (-2.4, 0.8)

-0.8 (-2.4, 0.8)

smokstatStopped> 1y ago

-0.6 (-1.3, 0.2)

-0.6 (-1.3, 0.2)

smokstatStopped in past year

-0.5 (-2.6, 1.6)

-0.5 (-2.6, 1.6)

smokstatSmoker

-1.1 (-1.8, -0.4)\*\*

-1.1 (-1.8, -0.4)\*\*

favdrinkCider

-1.1 (-2.5, 0.2)

-1.1 (-2.5, 0.2)

favdrinkMixed spirits

1.3 (-0.1, 2.6)

1.3 (-0.1, 2.6)

favdrinkOther

3.5 (-0.1, 7.0)

3.5 (-0.1, 7.0)

favdrinkSpirits alone

1.0 (-0.1, 2.1)

1.0 (-0.1, 2.1)

favdrinkWine

0.6 (-0.1, 1.4)

0.6 (-0.1, 1.4)

tryalclyc2Attempt to cut down in last 12 months

-0.5 (-1.2, 0.1)

-0.5 (-1.2, 0.1)

sexMen:f\_ageg16-24 years

-0.1 (-2.2, 2.0)

-0.1 (-2.2, 2.0)

sexMen:f\_ageg35-44 years

-0.1 (-2.3, 2.1)

-0.1 (-2.3, 2.1)

sexMen:f\_ageg45-54 years

-0.9 (-3.0, 1.2)

-0.9 (-3.0, 1.2)

sexMen:f\_ageg55-64 years

-1.1 (-3.3, 1.0)

-1.1 (-3.3, 1.0)

sexMen:f\_ageg65-74 years

1.3 (-1.1, 3.7)

1.3 (-1.1, 3.7)

sexMen:f\_ageg75+ years

-1.5 (-5.3, 2.3)

-1.5 (-5.3, 2.3)

Observations

9,850

9,850

R2

0.02

0.02

Adjusted R2

0.01

0.01

Residual Std. Error (df = 9816)

14.4

14.4

F Statistic (df = 33; 9816)

4.9\*\*\*

4.9\*\*\*

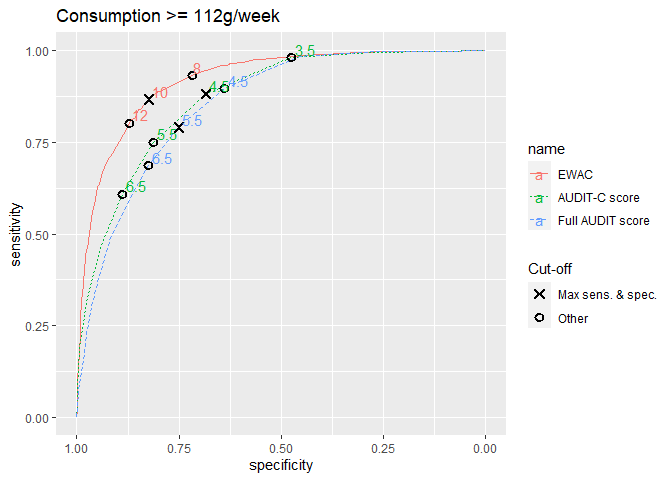
Notes:

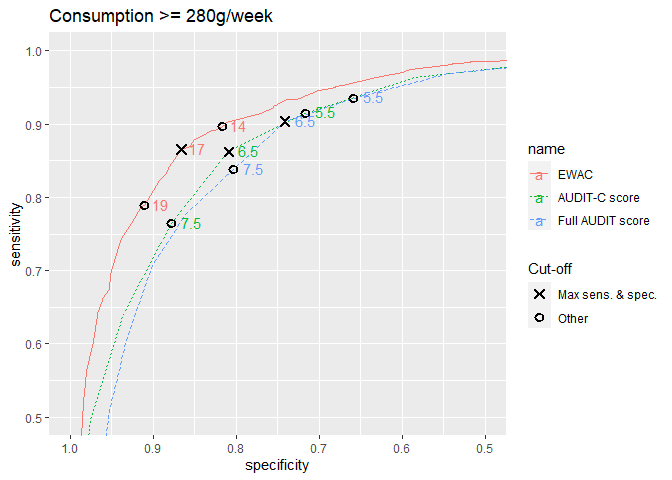
\*P < .05

\*\*P < .01

\*\*\*P < .001

## Supplementary information S6: Receiver operating characteristics curves





## Supplementary information S7: Demographic characteristics and alcohol consumption descriptives

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | <14 (N=17402) | 14-34 (N=3932) | 35-50 (N=582) | 50-400 (N=483) | Missing (N=4017) | Never drinks (AUDIT-1 = 0) (N=14399) | Total (N=40815) |
| **Gender** |  |  |  |  |  |  |  |
| Women | 8713 (50.1%) | 1200 (30.5%) | 124 (21.3%) | 104 (21.5%) | 2200 (54.8%) | 7996 (55.5%) | 20337 (49.8%) |
| Men | 8689 (49.9%) | 2732 (69.5%) | 458 (78.7%) | 379 (78.5%) | 1817 (45.2%) | 6403 (44.5%) | 20478 (50.2%) |
| In another way | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) |
| **Age** |  |  |  |  |  |  |  |
| N-Miss | 69 | 10 | 1 | 2 | 51 | 154 | 287 |
| 16-24 years | 2246 (13.0%) | 491 (12.5%) | 72 (12.4%) | 56 (11.6%) | 671 (16.9%) | 2383 (16.7%) | 5919 (14.6%) |
| 25-34 years | 2320 (13.4%) | 394 (10.0%) | 57 (9.8%) | 43 (8.9%) | 582 (14.7%) | 2400 (16.8%) | 5796 (14.3%) |
| 35-44 years | 2378 (13.7%) | 463 (11.8%) | 57 (9.8%) | 71 (14.8%) | 514 (13.0%) | 2243 (15.7%) | 5726 (14.1%) |
| 45-54 years | 2852 (16.5%) | 678 (17.3%) | 124 (21.3%) | 95 (19.8%) | 529 (13.3%) | 1845 (13.0%) | 6123 (15.1%) |
| 55-64 years | 2946 (17.0%) | 792 (20.2%) | 104 (17.9%) | 122 (25.4%) | 570 (14.4%) | 1808 (12.7%) | 6342 (15.6%) |
| 65-74 years | 2831 (16.3%) | 767 (19.6%) | 132 (22.7%) | 73 (15.2%) | 622 (15.7%) | 1828 (12.8%) | 6253 (15.4%) |
| 75+ years | 1760 (10.2%) | 337 (8.6%) | 35 (6.0%) | 21 (4.4%) | 478 (12.1%) | 1738 (12.2%) | 4369 (10.8%) |
| **Ethnic group** |  |  |  |  |  |  |  |
| N-Miss | 64 | 10 | 0 | 2 | 14 | 92 | 182 |
| White British | 15080 (87.0%) | 3673 (93.7%) | 552 (94.8%) | 445 (92.5%) | 3310 (82.7%) | 8637 (60.4%) | 31697 (78.0%) |
| White Other | 1147 (6.6%) | 133 (3.4%) | 15 (2.6%) | 18 (3.7%) | 327 (8.2%) | 1166 (8.1%) | 2806 (6.9%) |
| Mixed | 226 (1.3%) | 47 (1.2%) | 7 (1.2%) | 7 (1.5%) | 53 (1.3%) | 217 (1.5%) | 557 (1.4%) |
| Asian | 385 (2.2%) | 27 (0.7%) | 4 (0.7%) | 7 (1.5%) | 135 (3.4%) | 2901 (20.3%) | 3459 (8.5%) |
| Black | 372 (2.1%) | 28 (0.7%) | 2 (0.3%) | 2 (0.4%) | 146 (3.6%) | 1065 (7.4%) | 1615 (4.0%) |
| Other | 128 (0.7%) | 14 (0.4%) | 2 (0.3%) | 2 (0.4%) | 32 (0.8%) | 321 (2.2%) | 499 (1.2%) |
| **Highest qualification** |  |  |  |  |  |  |  |
| N-Miss | 71 | 7 | 2 | 1 | 30 | 109 | 220 |
| No qualification | 1692 (9.8%) | 349 (8.9%) | 57 (9.8%) | 69 (14.3%) | 681 (17.1%) | 3203 (22.4%) | 6051 (14.9%) |
| NVQ <= 3 | 8086 (46.7%) | 1796 (45.8%) | 284 (49.0%) | 247 (51.2%) | 2020 (50.7%) | 6677 (46.7%) | 19110 (47.1%) |
| NVQ4+ (degree) | 6405 (37.0%) | 1504 (38.3%) | 199 (34.3%) | 135 (28.0%) | 952 (23.9%) | 3238 (22.7%) | 12433 (30.6%) |
| Other | 1148 (6.6%) | 276 (7.0%) | 40 (6.9%) | 31 (6.4%) | 334 (8.4%) | 1172 (8.2%) | 3001 (7.4%) |
| **AUDIT-C score** |  |  |  |  |  |  |  |
| N-Miss | 25 | 2 | 2 | 2 | 143 | 0 | 174 |
| [0,1) | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 14399 (100.0%) | 14399 (35.4%) |
| [1,5) | 11895 (68.5%) | 550 (14.0%) | 19 (3.3%) | 19 (4.0%) | 3106 (80.2%) | 0 (0.0%) | 15589 (38.4%) |
| [5,8) | 4418 (25.4%) | 1856 (47.2%) | 136 (23.4%) | 76 (15.8%) | 578 (14.9%) | 0 (0.0%) | 7064 (17.4%) |
| [8,12] | 1064 (6.1%) | 1524 (38.8%) | 425 (73.3%) | 386 (80.2%) | 190 (4.9%) | 0 (0.0%) | 3589 (8.8%) |
| **Full AUDIT score** |  |  |  |  |  |  |  |
| N-Miss | 29 | 8 | 3 | 6 | 150 | 5 | 201 |
| [0,1) | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 14050 (97.6%) | 14050 (34.6%) |
| [1,8) | 15221 (87.6%) | 1897 (48.3%) | 104 (18.0%) | 67 (14.0%) | 3489 (90.2%) | 340 (2.4%) | 21118 (52.0%) |
| [8,16) | 2036 (11.7%) | 1846 (47.0%) | 403 (69.6%) | 239 (50.1%) | 322 (8.3%) | 4 (0.0%) | 4850 (11.9%) |
| [16,20) | 93 (0.5%) | 129 (3.3%) | 51 (8.8%) | 73 (15.3%) | 23 (0.6%) | 0 (0.0%) | 369 (0.9%) |
| [20,40] | 23 (0.1%) | 52 (1.3%) | 21 (3.6%) | 98 (20.5%) | 33 (0.9%) | 0 (0.0%) | 227 (0.6%) |
| **Favourite drink** |  |  |  |  |  |  |  |
| N-Miss | 11782 | 529 | 20 | 20 | 3208 | 14395 | 29954 |
| Beer | 2218 (39.5%) | 1493 (43.9%) | 259 (46.1%) | 228 (49.2%) | 334 (41.3%) | 1 (25.0%) | 4533 (41.7%) |
| Cider | 318 (5.7%) | 119 (3.5%) | 30 (5.3%) | 42 (9.1%) | 62 (7.7%) | 0 (0.0%) | 571 (5.3%) |
| Mixed spirits | 402 (7.2%) | 165 (4.8%) | 14 (2.5%) | 21 (4.5%) | 71 (8.8%) | 0 (0.0%) | 673 (6.2%) |
| Other | 57 (1.0%) | 12 (0.4%) | 4 (0.7%) | 1 (0.2%) | 16 (2.0%) | 2 (50.0%) | 92 (0.8%) |
| Spirits alone | 473 (8.4%) | 227 (6.7%) | 44 (7.8%) | 49 (10.6%) | 150 (18.5%) | 0 (0.0%) | 943 (8.7%) |
| Wine | 2152 (38.3%) | 1387 (40.8%) | 211 (37.5%) | 122 (26.3%) | 176 (21.8%) | 1 (25.0%) | 4049 (37.3%) |
| **Attempt to cut down** |  |  |  |  |  |  |  |
| N-Miss | 11781 | 529 | 20 | 19 | 3204 | 14395 | 29948 |
| No attempt to cut down | 4302 (76.5%) | 2453 (72.1%) | 400 (71.2%) | 316 (68.1%) | 642 (79.0%) | 2 (50.0%) | 8115 (74.7%) |
| Attempt to cut down in last 12 months | 1319 (23.5%) | 950 (27.9%) | 162 (28.8%) | 148 (31.9%) | 171 (21.0%) | 2 (50.0%) | 2752 (25.3%) |

## Supplementary information S8: STARD Checklist

# References

[1] Shield K, Manthey J, Rylett M, Probst C, Wettlaufer A, Parry CDH, et al. National, regional, and global burdens of disease from 2000 to 2016 attributable to alcohol use: a comparative risk assessment study. The Lancet Public Health 2020;5:e51–61. doi:[10.1016/S2468-2667(19)30231-2](https://doi.org/10.1016/S2468-2667(19)30231-2).

[2] American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. Washington, DC: American Psychiatric Association; 2013. doi:[10.1176/appi.books.9780890425596](https://doi.org/10.1176/appi.books.9780890425596).

[3] Health NI of, Excellence C. Alcohol-use disorders: prevention. Public health guideline [PH24]. 2010.

[4] National Institute of Health and Care Excellence. Alcohol-use disorders: diagnosis, assessment and management of harmful drinking and alcohol dependence. Clinical guideline [CG115]. 2011.

[5] Saunders JB, Degenhardt L, Reed GM, Poznyak V. Alcohol Use Disorders in ICD‐11: Past, Present, and Future. Alcoholism: Clinical and Experimental Research 2019;43:1617–31. doi:[10.1111/acer.14128](https://doi.org/10.1111/acer.14128).

[6] Babor TF, Higgins-Biddle JC, Saunders JB, Monteiro MG. The Alcohol Use Disorders Identification Test. Guidelines for Use in Primary Care. Second Edition. Geneva: World Health Organisation, Department of Mental Health; Substance Dependence; 2001.

[7] Bush K. The AUDIT Alcohol Consumption Questions (AUDIT-C): An Effective Brief Screening Test for Problem Drinking. Archives of Internal Medicine 1998;158:1789. doi:[10.1001/archinte.158.16.1789](https://doi.org/10.1001/archinte.158.16.1789).

[8] Rubinsky AD, Dawson DA, Williams EC, Kivlahan DR, Bradley KA. 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 2013;37:1380–90. doi:[10.1111/acer.12092](https://doi.org/10.1111/acer.12092).

[9] Meneses-Gaya C de, Zuardi AW, Loureiro SR, Crippa JAS. Alcohol Use Disorders Identification Test (AUDIT): An updated systematic review of psychometric properties. Psychology & Neuroscience 2009;2:83–97. doi:[10.3922/j.psns.2009.1.12](https://doi.org/10.3922/j.psns.2009.1.12).

[10] Nadkarni A, Garber A, Costa S, Wood S, Kumar S, MacKinnon N, et al. Auditing the AUDIT: A systematic review of cut-off scores for the Alcohol Use Disorders Identification Test (AUDIT) in low- and middle-income countries. Drug and Alcohol Dependence 2019;202:123–33. doi:[10.1016/j.drugalcdep.2019.04.031](https://doi.org/10.1016/j.drugalcdep.2019.04.031).

[11] Lavoie D. Alcohol identification and brief advice in England: A major plank in alcohol harm reduction policy. Drug and Alcohol Review 2010;29:608–11. doi:[10.1111/j.1465-3362.2010.00224.x](https://doi.org/10.1111/j.1465-3362.2010.00224.x).

[12] Corrao G, Bagnardi V, Zambon A, La Vecchia C. A meta-analysis of alcohol consumption and the risk of 15 diseases. Preventive Medicine 2004;38:613–9. doi:[10.1016/j.ypmed.2003.11.027](https://doi.org/10.1016/j.ypmed.2003.11.027).

[13] Johnson M, Jackson R, Guillaume L, Meier P, Goyder E. Barriers and facilitators to implementing screening and brief intervention for alcohol misuse: a systematic review of qualitative evidence. Journal of Public Health (Oxford, England) 2011;33:412–21. doi:[10.1093/pubmed/fdq095](https://doi.org/10.1093/pubmed/fdq095).

[14] Hutchings D, Cassidy P, Dallolio E, Pearson P, Heather N, Kaner E. Implementing screening and brief alcohol interventions in primary care: Views from both sides of the consultation. Primary Health Care Research and Development 2006;7:221–9. doi:[10.1191/1463423606pc292oa](https://doi.org/10.1191/1463423606pc292oa).

[15] McCormick KA, Cochran NE, Back AL, Merrill JO, Williams EC, Bradley KA. How Primary Care Providers Talk to Patients About Alcohol. Journal of General Internal Medicine 2006:060721075157048. doi:[10.1111/j.1525-1497.2006.00490.x](https://doi.org/10.1111/j.1525-1497.2006.00490.x).

[16] Beyer F, Lynch E, Kaner E. Brief Interventions in Primary Care: an Evidence Overview of Practitioner and Digital Intervention Programmes. Current Addiction Reports 2018;5:265–73. doi:[10.1007/s40429-018-0198-7](https://doi.org/10.1007/s40429-018-0198-7).

[17] Kaner E, Bland M, Cassidy P, Coulton S, Dale V, Deluca P, et al. Effectiveness of screening and brief alcohol intervention in primary care (SIPS trial): pragmatic cluster randomised controlled trial. BMJ 2013;346:e8501–1. doi:[10.1136/bmj.e8501](https://doi.org/10.1136/bmj.e8501).

[18] Crane D, Garnett C, Michie S, West R, Brown J. A smartphone app to reduce excessive alcohol consumption: Identifying the effectiveness of intervention components in a factorial randomised control trial. Scientific Reports 2018;8:4384. doi:[10.1038/s41598-018-22420-8](https://doi.org/10.1038/s41598-018-22420-8).

[19] Beard E, Brown J, West R, Acton C, Brennan A, Drummond C, et al. Protocol for a national monthly survey of alcohol use in England with 6-month follow-up: ’The Alcohol Toolkit Study’ Health behavior, health promotion and society. BMC Public Health 2015;15. doi:[10.1186/s12889-015-1542-7](https://doi.org/10.1186/s12889-015-1542-7).

[20] Kaner EF, Beyer FR, Garnett C, Crane D, Brown J, Muirhead C, et al. Personalised digital interventions for reducing hazardous and harmful alcohol consumption in community-dwelling populations. Cochrane Database of Systematic Reviews 2017. doi:[10.1002/14651858.CD011479.pub2](https://doi.org/10.1002/14651858.CD011479.pub2).

[21] Michie S, Whittington C, Hamoudi Z, Zarnani F, Tober G, West R. Identification of behaviour change techniques to reduce excessive alcohol consumption. Addiction 2012;107:1431–40. doi:[10.1111/j.1360-0443.2012.03845.x](https://doi.org/10.1111/j.1360-0443.2012.03845.x).

[22] Vinson DC. Alcohol is not a dichotomous variable. The Journal of Family Practice 1997;44:147–9.

[23] Rehm J, Marmet S, Anderson P, Gual A, Kraus L, Nutt DJ, et al. Defining Substance Use Disorders: Do We Really Need More Than Heavy Use? Alcohol and Alcoholism 2013;48:633–40. doi:[10.1093/alcalc/agt127](https://doi.org/10.1093/alcalc/agt127).

[24] Nutt DJ, Rehm J. Doing it by numbers: A simple approach to reducing the harms of alcohol. Journal of Psychopharmacology 2014;28:3–7. doi:[10.1177/0269881113512038](https://doi.org/10.1177/0269881113512038).

[25] Lemmens P, Tan ES, Knibbe R a. Measuring quantity and frequency of drinking in a general population survey: a comparison of five indices. Journal of Studies on Alcohol 1992;53:476–86. doi:[10.15288/jsa.1992.53.476](https://doi.org/10.15288/jsa.1992.53.476).

[26] Greenfield TK. Ways of measuring drinking patterns and the difference they make: experience with graduated frequencies. Journal of Substance Abuse 2000;12:33–49. doi:[10.1016/S0899-3289(00)00039-0](https://doi.org/10.1016/S0899-3289(00)00039-0).

[27] Livingston M, Callinan S. Underreporting in Alcohol Surveys: Whose Drinking Is Underestimated? Journal of Studies on Alcohol and Drugs 2015;76:158–64. doi:[10.15288/jsad.2015.76.158](https://doi.org/10.15288/jsad.2015.76.158).

[28] Heeb J-L, Gmel G. Measuring alcohol consumption: A comparison of graduated frequency, quantity frequency, and weekly recall diary methods in a general population survey. Addictive Behaviors 2005;30:403–13. doi:[10.1016/j.addbeh.2004.04.022](https://doi.org/10.1016/j.addbeh.2004.04.022).

[29] Rehm J. Measuring Quantity, Frequency, and Volume of Drinking. Alcoholism: Clinical and Experimental Research 1998;22:4s–14s. doi:[10.1111/j.1530-0277.1998.tb04368.x](https://doi.org/10.1111/j.1530-0277.1998.tb04368.x).

[30] Stockwell T, Donath S, Cooper-Stanbury M, Chikritzhs T, Catalano P, Mateo C. Under-reporting of alcohol consumption in household surveys: A comparison of quantity-frequency, graduated-frequency and recent recall. Addiction 2004;99:1024–33. doi:[10.1111/j.1360-0443.2004.00815.x](https://doi.org/10.1111/j.1360-0443.2004.00815.x).

[31] Stockwell T, Zhao J, Macdonald S. Who under-reports their alcohol consumption in telephone surveys and by how much? An application of the ’yesterday method’ in a national Canadian substance use survey. Addiction (Abingdon, England) 2014;109:1657–66. doi:[10.1111/add.12609](https://doi.org/10.1111/add.12609).

[32] Stockwell T, Zhao J, Greenfield T, Li J, Livingston M, Meng Y. Estimating under- and over-reporting of drinking in national surveys of alcohol consumption: identification of consistent biases across four English-speaking countries. Addiction 2016;111:1203–13. doi:[10.1111/add.13373](https://doi.org/10.1111/add.13373).

[33] Public Health England. Alcohol sales in England in 2014: Analysis to assess suitability for inclusion as an indicator in the Local Alcohol Profiles for England. 2017.

[34] NatCen Social Research, Royal Free and University College Medical School. Health Survey for England, 2011 [computer file]. Colchester, Essex: UK Data Archive [distributor], April 2013. SN: 7260. Colchester, Essex: 2013. doi:[10.5255/UKDA-SN-7260-1](https://doi.org/10.5255/UKDA-SN-7260-1).

[35] Boniface S, Kneale J, Shelton N. Drinking pattern is more strongly associated with under-reporting of alcohol consumption than socio-demographic factors: evidence from a mixed-methods study. BMC Public Health 2014;14:1297. doi:[10.1186/1471-2458-14-1297](https://doi.org/10.1186/1471-2458-14-1297).

[36] Dutey-Magni P, Sinclair J, Brown J. Concurrent validity of an Estimator of Weekly Alcohol Consumption (EWAC) based on the Extended AUDIT 2018. doi:[10.17605/OSF.IO/7WE4M](https://doi.org/10.17605/OSF.IO/7WE4M).

[37] R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2017.

[38] Wickham H. Tidyverse: Easily install and load the ’tidyverse’ 2017.

[39] Stan Development Team. RStan: The R interface to Stan 2018.

[40] Dutey-Magni P. Concurrent validity of an Estimator of Weekly Alcohol Consumption (EWAC) based on the Extended AUDIT 2020.

[41] DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. Biometrics 1988;44:837. doi:[10.2307/2531595](https://doi.org/10.2307/2531595).

[42] Frank D, DeBenedetti AF, Volk RJ, Williams EC, Kivlahan DR, Bradley KA. Effectiveness of the AUDIT-C as a Screening Test for Alcohol Misuse in Three Race/Ethnic Groups. Journal of General Internal Medicine 2008;23:781–7. doi:[10.1007/s11606-008-0594-0](https://doi.org/10.1007/s11606-008-0594-0).

[43] Bradley KA, McDonell MB, Bush K, Kivlahan DR, Diehr P, Fihn SD. The AUDIT Alcohol Consumption Questions: Reliability, Validity, and Responsiveness to Change in Older Male Primary Care Patients. Alcoholism: Clinical and Experimental Research 1998;22:1842–9. doi:[10.1111/j.1530-0277.1998.tb03991.x](https://doi.org/10.1111/j.1530-0277.1998.tb03991.x).

[44] Jellinek EM. The disease concept of alcoholism. New Haven: Hillhouse Press; 1960. doi:[10.1037/14090-000](https://doi.org/10.1037/14090-000).

[45] Leggio L, Kenna GA, Fenton M, Bonenfant E, Swift RM. Typologies of Alcohol Dependence. From Jellinek to Genetics and Beyond. Neuropsychology Review 2009;19:115–29. doi:[10.1007/s11065-008-9080-z](https://doi.org/10.1007/s11065-008-9080-z).

[46] Gorman E, Leyland AH, McCartney G, White IR, Katikireddi SV, Rutherford L, et al. Assessing the Representativeness of Population-Sampled Health Surveys Through Linkage to Administrative Data on Alcohol-Related Outcomes. American Journal of Epidemiology 2014;180:941–8. doi:[10.1093/aje/kwu207](https://doi.org/10.1093/aje/kwu207).

[47] Christensen AI, Ekholm O, Gray L, Glümer C, Juel K. What is wrong with non-respondents? Alcohol-, drug- and smoking-related mortality and morbidity in a 12-year follow-up study of respondents and non-respondents in the Danish Health and Morbidity Survey. Addiction 2015;110:1505–12. doi:[10.1111/add.12939](https://doi.org/10.1111/add.12939).

[48] Boniface S, Scholes S, Shelton N, Connor J. Assessment of Non-Response Bias in Estimates of Alcohol Consumption: Applying the Continuum of Resistance Model in a General Population Survey in England. PLOS ONE 2017;12:e0170892. doi:[10.1371/journal.pone.0170892](https://doi.org/10.1371/journal.pone.0170892).

[49] Vocht F de, Brown J, Beard E, Angus C, Brennan A, Michie S, et al. Temporal patterns of alcohol consumption and attempts to reduce alcohol intake in England. BMC Public Health 2016;16:917. doi:[10.1186/s12889-016-3542-7](https://doi.org/10.1186/s12889-016-3542-7).

[50] Greenfield TK, Kerr WC, Bond J, Ye Y, Stockwell T. Improving Graduated Frequencies Alcohol Measures for Monitoring Consumption Patterns: Results from an Australian National Survey and a US Diary Validity Study. Contemporary Drug Problems 2009;36:705–33. doi:[10.1177/009145090903600320](https://doi.org/10.1177/009145090903600320).

[51] Hilton ME. A comparison of a prospective diary and two summary recall techniques for recording alcohol consumption. British Journal of Addiction 1989;84:1085–92. doi:[10.1111/j.1360-0443.1989.tb00792.x](https://doi.org/10.1111/j.1360-0443.1989.tb00792.x).

[52] UK Health and Social Care Information Centre. UK SNOMED CT Drug Extension, RF2: Full, Snapshot & Delta 2018.

[53] Haroon S, Wooldridge D, Hoogewerf J, Mittal A, Bhala N, O’Donnell A, et al. Information standards for recording alcohol use in electronic health records: Project report 2018.

[54] Mongan D, Long J. Standard drink measures throughout Europe; peoples’ understanding of standard drinks. RARHA: Joint Actional on Reducing Alcohol Related Harm; 2015.

[55] Kalinowski A, Humphreys K. Governmental standard drink definitions and low-risk alcohol consumption guidelines in 37 countries. Addiction 2016;111:1293–8. doi:[10.1111/add.13341](https://doi.org/10.1111/add.13341).

[56] UK Office for National Statistics. Drinking: adults’ behaviour and knowledge in 2008. Opinions (Omnibus) Survey Report No. 39. 2009.

[57] Brown J, West R, Angus C, Beard E, Brennan A, Drummond C, et al. Comparison of brief interventions in primary care on smoking and excessive alcohol consumption: A population survey in england. British Journal of General Practice 2016;66:e1–9. doi:[10.3399/bjgp16X683149](https://doi.org/10.3399/bjgp16X683149).

[58] Rehm J, Anderson P, Manthey J, Shield KD, Struzzo P, Wojnar M, et al. Alcohol Use Disorders in Primary Health Care: What Do We Know and Where Do We Go? Alcohol and Alcoholism 2016;51:422–7. doi:[10.1093/alcalc/agv127](https://doi.org/10.1093/alcalc/agv127).