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

Table of Contents

# Abstract

**Background and Aims:** To develop and validate an Estimator of Weekly Alcohol Consumption (EWAC) using information collected by the consumption questions of the Extended Alcohol Use Disorders Identification Tool (Extended AUDIT-C). The Extended AUDIT-C contains a greater range of response items to quantify drinking frequency and quantity than the AUDIT-C.

**Methods:** *Study 1* evaluates the concurrent validity of EWAC against Graduated-Frequency (GF) measurements in a sample of 22,404 English household residents, testing whether (a) the mean deviation (metric of bias) exceeded a target of 1 UK unit, and (b) the root mean squared deviation (metric of total error) exceeded 2 UK units. Receiver Operating Characteristics in predicting alcohol consumption in excess of 112g/week (14 UK units) and 280g/week (35 UK units) are compared in EWAC, AUDIT-C, and the full AUDIT. *Study 2* evaluates the concurrent validity of EWAC against 28-day Timeline Followback (TLFB) measurements in a sample of 98 outpatient clinic attendees in one UK hospital, and tests the administration mode effect on bias (self- vs researcher-administered). *Study 3* compares the EWAC’s empirical distribution of alcohol consumption in England with other sources of data.

**Results:** The EWAC estimates usual alcohol consumption with a positive bias of 0.2 UK units [95% confidence interval: 0.04, 0.3] and a total error of 11 UK units/week [9.8, 12.0] when compared to GF, or 5 UK units/week [3.9, 7.6] when compared to TLFB. No statistical evidence of an effect of self-administration on bias ( = 0.399) was found. The EWAC predicts alcohol consumption in excess of 112g/week with a significantly greater area under the curve (0.921 [0.917, 0.925]) than the AUDIT-C (0.871 [0.866, 0.876]) or the full AUDIT (0.854 [0.849, 0.860]). Similarly, the EWAC predicts consumption in excess of 280g/week with a significantly greater area under the curve (0.936 [0.929; 0.943]) than the AUDIT-C (0.913 [0.904, 0.921]) or the full AUDIT (0.901 [0.893; 0.909]). *Studies 1-3* suggest that the EWAC may underestimate the alcohol consumption of higher-risk drinkers based on comparisons with other data sources.

**Conclusions:** The EWAC provides greater screening specificity than the AUDIT-C or the full AUDIT. In addition, it can provide additional personalised information to patients in a more understandable and potentially less stigmatising way than AUDIT scores or risk categorisations. The EWAC has potential to generate personalised feedback within digital alcohol 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.

Clinical guidelines aim to prevent [[2](#ref-NICE-PH24)], treat and reduce [[3](#ref-NICE-CG115)] harm from alcohol consumption. Most recommend systematic screening for alcohol use disorders using a validated diagnostic tool. Recent clinical research has advocated a greater focus on monitoring alcohol consumption in the same way as other health risk factors (eg body mass index, blood pressure or cholesterol), keeping individuals regularly informed of their current intake and the associated risks [[4](#ref-Nutt2014),[5](#ref-Rehm2016)]. This approach is intended to facilitate timely lifestyle interventions to prevent the escalation to serious levels of alcohol-related harm. Simultaneously, public health recommendations [[6](#ref-AU.alcoguidelines2009),[7](#ref-AlcoholCMO2016b)] now promote the overall reduction of alcohol consumption, since there is absolute ‘safe level’ at which alcohol does not increase general health risk [[8](#ref-Wood2018)].

Assessment of alcohol consumption is not well embedded in clinical practice. Reasons include: difficulty in the self-assessment of alcohol consumption, lack of a single biological marker, time pressures, and stigma. This creates 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.

The objective of this paper is to develop and validate a fast and easy-to-complete Estimator of Weekly Alcohol Consumption (EWAC) by repurposing an existing screening tool. In many countries, the most widely used screening tool is the Alcohol Use Disorders Identification Test (AUDIT) [[9](#ref-Babor2001)], a 10-item questionnaire schedule employed in clinical practice and research as a diagnostic screening tool. The short, 3-item consumption questions known as AUDIT-C has good accuracy for both: (a) interview-based clinical diagnoses; and (b) consumption in excess of maximum recommended intakes (eg 140g/week in Australia, 112 g/week in UK) in a variety of populations [[10](#ref-DeMeneses-Gaya2009)].

The present study aims to validate the EWAC against continuous measures of alcohol consumption, as opposed to against a particular clinical diagnosis or consumption threshold. To this end, we adopt a variant of the AUDIT-C schedule known as the ‘Extended AUDIT-C’, which improves the granularity of the information collected with a greater range of response options on quantity and frequency (AUDIT items 1 and 2). The Extended AUDIT-C has been used in UK research as part of two trials [[11](#ref-Kaner2013c),[12](#ref-Crane2018)] and one continuous household survey [[13](#ref-Beard2015a)] to capture greater information on the higher risk drinkers, based on the observation that AUDIT consumption response items are right-truncated.

This paper is made up of three investigations. Study 1 estimates coefficients to apply to each of the AUDIT-C response items to compute an EWAC using data from a large English household survey, the Alcohol Toolkit Study (ATS). It then tests the concurrent validity of the EWAC against Graduated-Frequency (GF) measurements across a number of demographic subpopulations in England. Study 2 tests the concurrent validity of the EWAC against the 28-day Timeline Followback (TLFB) in a clinical population (hospital clinic outpatients). Study 3 compares four sources of data (the EWAC, the ATS GF schedule, the Health Survey for England, and alcohol sales official statistics) on the population-wide cumulative distribution of alcohol consumption in England.

# Methods

## Approach

This paper develops and validates an Estimator of usual Weekly Alcohol Consumption in 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 item 1) and quantity of drinking (AUDIT item 2) can be used to estimate usual alcohol consumption, with adjustment for occasional heavy use (AUDIT item 3), following methods developed for quantity-frequency-variability instruments [[14](#ref-Lemmens1992)].

In practice, for every individual , the EWAC is computed as the product of and (AUDIT items 1 and 2 respectively) adjusted with the frequency of intense drinking (AUDIT item 3):

where denotes the average number of 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 item 2
* a statistical model to estimate the coefficients empirically from alcohol consumption data

All results are reported in UK alcohol units (8g or 10mL of pure alcohol). Analyses are conducted in R [[15](#ref-RCoreTeam2017)] using packages tidyverse and rstan [[16](#ref-package-tidyverse),[17](#ref-package-rstan)]. Computer scripts for all analyses are available online [[18](#ref-Dutey2020)].

## Data sources

A longstanding obstacle in alcohol research and treatment is that there is no undisputed ‘gold standard’ or biomarker for assessing the level and duration of alcohol consumption. Instead, a number of instruments exist which measure self-reported alcohol consumption with varying validity and reliability over different time periods. A comprehensive review of these is summarised in Table 1 below. Prospective diaries tend to record higher alcohol consumption by minimising recall bias, followed by GF, while lower levels seem to be recorded with QF measures [[19](#ref-Heeb2005),[20](#ref-Rehm1998)].

Table 1: Alcohol schedules: selective comparison

Schedule

Bias

Variance

Measures variability

Graduated frequency (GF)

Unclear

Low

Yes

Quantity-Frequency (QF)

Unclear

Low

No

Quantity-Frequency-Variability (QFV)

Unclear

Low

Yes

Yesterday’ recall

Minimal

High

No

Timeline followback (TLFB)

Low

Low

Yes

Prospective diary

Low

Low

Yes

Table 2 summarises the four data sources used in this paper alongside references to methodological descriptions.

Table 2: Description of data sources

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Source | Time | Study population | Measurements | Validation | Sample size | Design |
| Alcohol Toolkit Study (ATS) [[13](#ref-Beard2015a)] | November 2015–October 2017 (waves 110–133) | Residents of private English households aged 16 years and over | Interviewer-led Extended AUDIT-C and GF | [[19](#ref-Heeb2005)], [[21](#ref-Greenfield2009)] | 40,832 | Cross-sectional sample survey |
| Health Survey for England (HSE) [[22](#ref-NatCenSocialResearch2013)] | 2011 | Residents of private English households aged 16 years and over | Interviewer-led beverage-specific QF | Consistent with a prospective diary [[23](#ref-Boniface2014)] and yesterday recall [[24](#ref-Stockwell2016)]. | 8,610 | Cross-sectional sample survey |
| Hospital questionnaire study (HOSP) | February 2019–October 2019 | Outpatient and day case hospital patients aged 18 years and over | Interviewer/self-administered Extended AUDIT-C and interviewer-led TLFB | TLFB records fewer drinking days and thus lower consumption than a prospective diary [[25](#ref-Grant1995)]. This recall bias increases with the number of days elapsed [[26](#ref-Hoeppner2010),[27](#ref-Vinson2003)] | 98 | Cross-sectional study with block randomisation to (a) self-administered AUDIT or (b) researcher-administered AUDIT. |
| Alcohol retail sales [[28](#ref-PHE2017)] | 2014 | English population aged 18 years and over | – | [[28](#ref-PHE2017)] | – | Ratio of all alcohol produced or processed in the UK, as well as alcohol imported into the UK for sale and consumption, over the mid-year population estimate |

## Study 1: EWAC estimation and concurrent validity

Study 1 evaluates the accuracy of the EWAC derived from a researcher-administered Extended AUDIT-C in community households. A pre-registered protocol for this study is available online [[29](#ref-Dutey2018)]. It uses the GF schedule from the ATS as the reference standard. Two sets of coefficients are compared: AUDIT response item interval midpoint; and coefficients estimated using a statistical model (supplementary material 1-2).

First, the agreement between the EWAC and the GF is quantified by studying two types of deviations:

* 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 gold standard. We test the hypothesis that the RMSD is greater than 2 UK units using a one-sided homogeneity test.

Second, we examine whether the EWAC’s validity varies across population subgroups.

* the simple deviation is regressed in a linear model to test for subgroup differences in MD
* the squared deviation $(\rm{EWAC}-\rm{GF})^2$ is regressed in a log-transformed linear model to test subgroup differences in the geometric mean squared deviation. Model coefficients are then 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 include the following predictors: sex by age group; ethnic group; highest educational qualification; religion; smoking status. Additional models are fitted solely in respondents with an AUDIT-C score of 5 or more or an AUDIT score of 8 or more (under the ICD terminology, hazardous and harmful alcohol use), using extra variables measures 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 (eg by drinking less, choosing lower strength alcohol or using smaller glasses).

Third, we test whether the EWAC is superior to the traditional AUDIT and AUDIT-C scores at predicting drinking in excess of 14 or 35 UK units per week. In the UK, an AUDIT-C score of 5-7 is often categorised as ‘increasing risk’, and scores of 8 or more are categorised as ‘higher risk’ drinking [[30](#ref-Lavoie2010)]. The study tests the hypothesis that the EWAC’s receiver operating characteristics’ full area under the curve (AUC) is greater that the comparators using nonparametric paired AUC tests [[31](#ref-Delong1988)].

The ATS data ( 40,832) is affected by missing data: 35% of respondents ( 14,408) reported never drinking alcohol in AUDIT item 1 and were not asked any further AUDIT or GF questions. These participants are excluded from the analysis. A further 4,020 respondents (0.2% of those reporting drinking in AUDIT item 1) did not have a valid GF alcohol consumption record and were also excluded. In total, 22,404 valid observations remain for the diagnostic analysis, in which missing GF data is assumed to be missing at random conditionally on the Extended AUDIT-C responses. In the subgroup analysis (Study 2 below), a further 530 repondents (0.5%) were assumed to have data missing at random and were excluded.

## Study 2: Concurrent validity in hospital outpatients

Study 2 aims to confirm the robustness of findings (a) in a clinical population; (b) when the Extended AUDIT-C is self-administered; (c) using a different reference standard: the 28-day TLFB. A pilot study recruited participants prospectively from a range of clinics (endoscopy day cases, outpatient orthopaedics, ophthalmology, rheumatology, young adults, and managed care clinics) at a large acute hospital in Southampton, UK, until a target minimum of 100 participants provided usable data on the Extended AUDIT-C. A total of 118 participants consented to participating, and block-randomised to one of two groups:

1. self-administered Extended AUDIT-C (n=59)
2. researcher-administered Extended AUDIT-C (n=59).

Once the Extended AUDIT-C was complete, both groups were administered a 28-day TFLB by the researcher.

A total of 16 participants (14%) did not complete the Extended AUDIT-C successfully and were excluded from the analysis. A further 4 participants who had not provided consumption estimates for at least 14 days were also excluded. Out of 98 participants included in the analysis, 89 completed all 28 days of the TLFB successfully. Just 1% (39/2744) of all daily measures were missing.

For each participant, the mean of TLFB daily measures was multiplied by 7 to compute a weekly average consumption and compared with the EWAC using identical tests as in Study 1. Next, we estimated the effect of the mode of administration (self vs researcher-administered) on MD, and tested the null hypothesis . With an RMSD of 6, a sample of n = 48 within each administration group is sufficient to detect an effect of at least 3.5 UK units (28g) with 80% power and 95% confidence using a two-sided independent groups *t*-test.

## Study 3: Aggregate concurrent validity

Study 3 examines the consistency in aggregate alcohol consumption estimates across England in residents aged 18 years and over. We plot the empirical cumulative distributions of alcohol consumption given by (1) the EWAC estimated from the ATS; (2) the quantity-frequency estimated in the ATS; (3) the beverage-specific estimators in HSE in 2011; (4) the prospective diary estimator in HSE 2011. In this analysis, survey weights are used: in (1-3), poststratification weights estimated using calibration and age-sex mid-year population estimates; in (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 for England from [[28](#ref-PHE2017)].

# Results

## Study 1: EWAC estimation and validation

### Overall bias and accuracy

The first step involved choosing a set of coefficients to compute the EWAC (supplementary materials 1–2).

The EWAC computed with the midpoint of the AUDIT item intervals had a Pearson’s correlation = 0.69 [95% CI: 0.69, 0.70]. It produced a mean deviation (MD) of 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 (statistical model reported in supplementary materials 2) provided a small improvement: = 0.71 [0.71, 0.72] (Kendall’s rank correlation = .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 3 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 (eg ‘10 to 14 units’) would contain the reference standard for half of individuals, while an interval estimate defined as the EWAC 3 units (eg ‘9 to 15 units’) would contain the reference standard for 60% of individuals.

Table 3: 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.

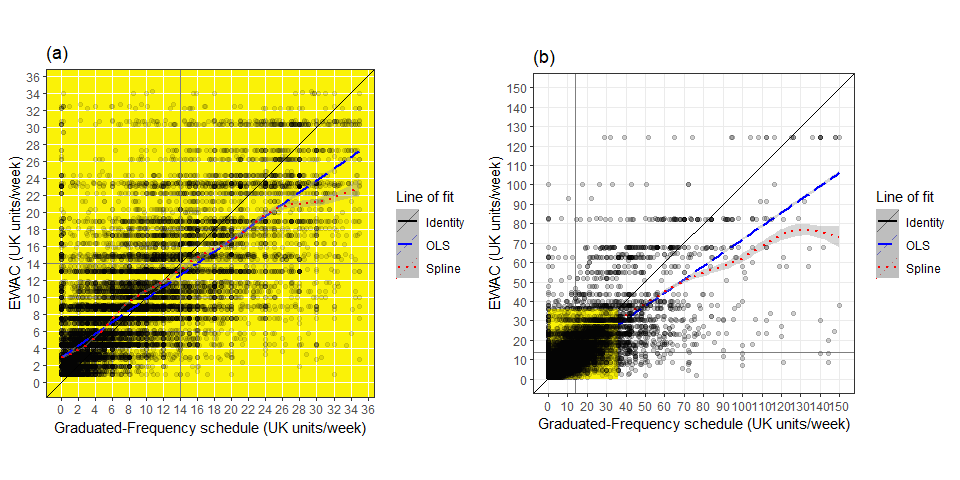


Figure 1: Plots of EWAC against GF in (a) low/increasing risk ATS respondents (n=21,338) and (b) all ATS respondents (n=22,373)

### Subgroup accuracy

Next, the MD and RMSD were regressed against respondent characteristics in order to identify subgroups with heterogeneous bias or precision (supplementary material 3, Table 8). The reference category’s key characteristics are: females aged between 25 and 34 years of White ethnicity without educational qualifications, who never smoked. The model predictors explain a very modest proportion of both MD and RMSD ( statistics < 2%). Nevertheless, specific subgroups do exhibit very different MD and RMSD.

Figure 7 summarises the MD for a selection of 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. Similar results were found in increasing risk drinkers, without significant evidence of an effect of favourite drink or attempts to reduce alcohol intake in the past year (supplementary material 3, Table 9)

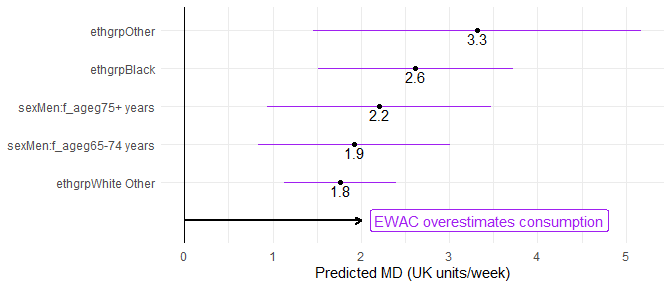


Figure 2: Forest plot of modelled MD for selected subgroups

Figure 3 shows those subgroups with an RMSD found to be 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 is 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 is also 44% greater [29; 60%] greater in men and 34% greater [19; 50%] in respondents aged 16 to 24 years. Conversely, error is 35 to 70% smaller in White Other, Black and Other ethnic groups.

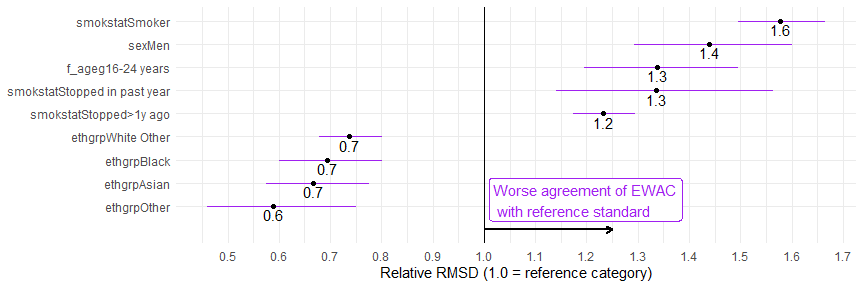


Figure 3: Forest plot of RMSD ratio (selected subgroups to reference category)

Figure 4 visualises the same analysis in increasing risk drinkers exclusively (AUDIT-C 5 or AUDIT 8). With a higher mean alcohol consumption, results differ 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 seem to significantly improve the agreement between EWAC and the reference standard. School and degree-level qualifications reduced RMSD by 24% [12; 37%] and 37% [23; 52%] smaller respectively with unchanged MD, 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.

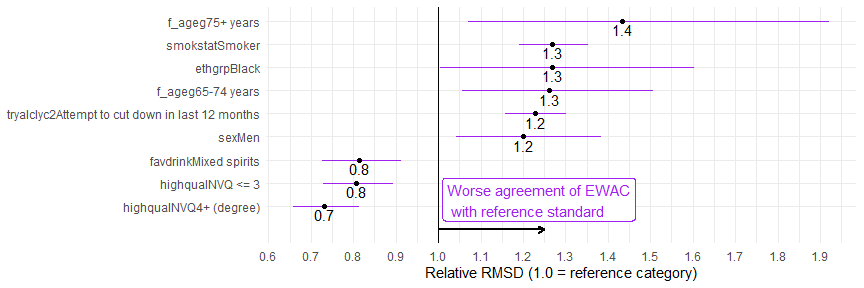


Figure 4: Forest plot of RMSD ratio (selected subgroups to reference category) in respondents with a hazardous/harmful alcohol use (AUDIT-C>=5 or AUDIT>=8

### Receiver Operating Characteristics

The last step of this analysis was to examine the EWAC’s ability to predict that participants’ consumption exceeds 14 or 35 UK units/week. Receiver operating characteristics curves (available in Supplementary Material 4) are summarised with respect to the full area under the curve as well as sensitivity/specificity for the best thresholds in Table 4 and Table 5.

**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 is 10 units/week. The sensitivity at this threshold is identical to AUDIT-C, but specificity is increased by 13 percentage points. Using the nominal cut-off of 14 units/week on the EWAC increases specificity to 0.922, at the cost of a smaller sensitivity of 0.705.

Table 4: 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

**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).

Table 5: 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.

## Study 2: Concurrent validity in hospital patients

A total of 59 participants (60%) met the definition of low-risk alcohol consumption (AUDIT-C score below 5), 16 participants (16%) increasing-risk alcohol consumption (AUDIT-C score 5 to 7), and 23 participants (23%) high-risk alcohol consumption (AUDIT-C score 8 or more).

MD was estimated at -0.7 unit/week [95%CI: -1.9, 0.5], which does not provide any evidence of bias greater than 1 unit/week ( = 0.270). As for error, RMSD was estimated at 6.0 [95%CI: 3.9, 7.6] and was statistically significantly greater than 2 units ( < 0.001). A potential reason for RMSD being considerably smaller than in Study 1 is the distribution of alcohol consumption of this small pilot dataset. Respondents’ alcohol intake as estimated by TLFB was low, with a mean of 7 units/week; a median of 2 units/week, and a 90th percentile of 23 units/week. Consequently, the probability of observing the strong deviations that can be observed in presence of high alcohol consumption data was low.

We investigated whether the mode of administration of the Extended AUDIT-C had an effect. In the self-administered group, MD = -1.2, compared with MD = -0.1. This suggests, at the first look, that EWAC was 1.0 units greater [-1.4; 3.4] when the Extended AUDIT-C is administered by a researcher, but this difference is not statistically significant ( = 0.399). A larger sample size would be required to address this question with greater precision.

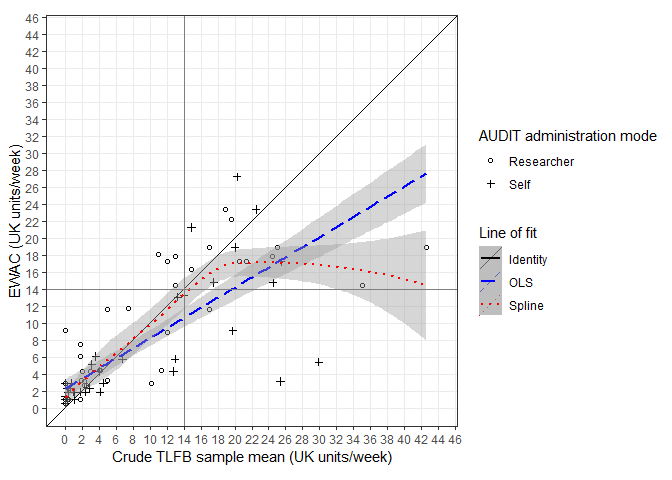


Figure 5: Plot of EWAC estimates against TLFB estimates in hospital participants with lines of fit

## Study 3: Empirical distribution functions

Table 6 provides estimates of total alcohol consumption in adults residing in private households 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 are just 71% of the HSE prospective diary, considered a very reliable estimate.

Table 6: 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 6 suggests that the EWAC, like the ATS GF, estimates a higher prevalence of low risk consumption ( 14 units/week) and increased risk consumption than HSE. In contrast very high alcohol consumption ( 50 units/week) is higher in HSE. This may be due to a combination of difference in sampling coverage, nonresponse bias, or measurement error in the alcohol schedules.

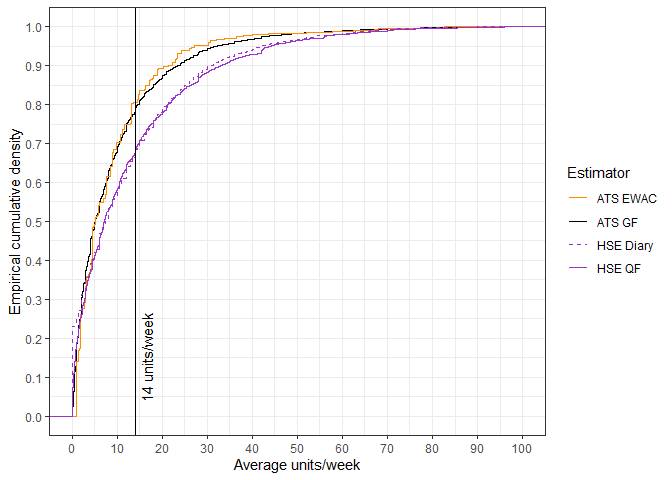


Figure 6: Empirical cumulative distribution function of weekly alcohol consumption in England according to four alcohol schedules in residents aged 18 years and over

# Discussion

## Main findings

This paper examined the predictive capabilities of the Extended AUDIT-C in assessing alcohol consumption in two English populations: community dwelling household residents, and hospital clinic outpatients. The Extended AUDIT-C is a variant of the AUDIT-C containing a choice of 6 response items to quantify drinking frequency, and 7 response items to quantify the average quantity consumed on any drinking day. The resulting EWAC estimates usual alcohol consumption with a mean precision of 11 units/week when compared to the GF alcohol schedule, or 5 units/week when compared to a 28-day TLFB in a sample of lower-risk alcohol drinkers. Although bias is mostly consistent across subgroups examined (age/sex, education, smoking status, religion), there is strong evidence that EWAC overestimates alcohol consumption by 2-3 units/week in Black and Other ethnic groups. At the same time, average precision is better in Black, Asian and Other ethnic groups. This is likely due to a lower mean alcohol consumption than White British and Mixed ethnic groups. We also noted a weaker precision for both current and ex-smokers. Using TLFB as a reference standard, similar results are found in a clinical population (Study 2), with no strong evidence that the EWAC is less valid when the Extended AUDIT-C is self-administered than when administered by a researcher.

## 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 [[32](#ref-Rubinsky2013)] previously reported mean consumption by AUDIT-C score, yet without quantifying bias or precision of such a measure. Other studies have studied the AUDIT-C’s potential in estimating alcohol consumption, but only in relation to predicting consumption in excess of a predefined threshold [[10](#ref-DeMeneses-Gaya2009)]. Such studies achieved AUCs ranging 0.83-0.96. In Study 1, we demonstrated the EWAC’s superiority for predicting GF 14 units/week. While its AUC of 0.921 [95% CI: 0.917; 0.925] is comparable with many other studies, the specificity gain from 0.684 to 0.876 places the EWAC with the other best-performing alcohol consumption estimators.

This study provides strong confidence in the internal and external validity of findings in England on account of the large sample size of the ATS, the range of subgroup analyses reported, and replication of the analysis in a clinical population using a second reference standard, the TLFB (Study 2).

We note some 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 conditions such as addiction or alcohol-related liver disease. The quality of alcohol survey data may also impact on the internal validity of Study 1.

The selected reference standard has a strong influence on estimates of precision (RMSD). The reference standard’s own error will inflate the RMSD if it does not correlate with the EWAC’s error. This means some of the RMSD is likely to be attributable to error in the reference standard rather than the EWAC. For example, a US telephone interview study on 119 participants [[21](#ref-Greenfield2009)] measured a Pearson’s correlation coefficient of = 0.86 and 0.87 between drinking frequencies and volumes measured by a GF interview and a subsequent 28 day prospective diary. A Kendall’s rank correlation = 0.41 between GF and a 7-day prospective diary was found in a telephone interview study of 773 participants across Switzerland [[19](#ref-Heeb2005)]. For reference, our study measured a correlation coefficient = 0.71 and = 0.63 between EWAC and GF, that is, a similar level of agreement as GF with TLFB. Since both reference standards employed in studies 1 and 2 can be safely assumed to be affected by substantial error of their own, our measures of RMSD can be regarded as very conservative estimators of precision. This means precision is likely to be better in reality than reported by studies 1 and 2.

Despite its large sample size, the ATS exhibits limitations of its own. 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 nonrespondents may have higher levels of alcohol consumption and harm [[33](#ref-Gorman2014)–[35](#ref-Boniface2017)]. Furthermore, the ATS potentially misses some (mostly low-risk) drinkers, even among respondents: the proportion of respondents classified as non-drinkers (based on AUDIT item 1) in ATS is 10% higher than in HSE data, in which respondents are asked to confirm whether they never drink alcohol, or only drink ‘very occasionally’ [[36](#ref-DeVocht2016)]. All these factors are likely to affect MD and RMSD estimates reported in Study 1, and explain some of the discrepancy with Study 3 findings.

## Potential applications

While retaining the AUDIT’s strengths (speed, accuracy, international standardisation), the Extended AUDIT improves the granularity of the information collected on alcohol consumption, and captures greater information from higher-risk drinkers, by remediating the right-truncation of the traditional AUDIT consumption items. The EWAC enhances the Extended AUDIT’s health education and promotion qualities, by translating it into a more understandable scale: alcohol consumption. This has the merit of focusing the recipient’s attention to alcohol units or grams, and can 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 [[30](#ref-Lavoie2010)]. Knowledge of alcohol beverage content is generally poor [[37](#ref-Mongan2015)], and many countries have not adopted a measure of standard drinks [[38](#ref-Kalinowski2016)]. In England, two thirds of drinkers could assess the standard drink equivalent in wine or beer of one alcohol unit [[39](#ref-ONS2009)].

A more clinical approach to alcohol harm prevention would be to treat alcohol consumption like any other risk factor: blood pressure or cholesterol are frequently monitored within health systems as part of standard care and can act as a trigger for behaviour change. In a similar way, [[4](#ref-Nutt2014)] and [[5](#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.

The proposed EWAC can fulfil the same alcohol use disorder screening functions as the AUDIT-C, while at the same time providing transparent and personalised information to patients in a less stigmatising way than terminology previously in use (eg 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 communication 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 lifestyle. The EWAC appears particularly suitable to generate personalised feedback and monitoring as part of digital interventions.

# Ethics

Study 1 and 3 were approved by the University of Southampton’s Faculty of Medicine Ethics Committee (ERGO 44682). Study 2 was approved by the Health Research Authority National Research Ethics Service (IRAS 247458; REC 18/SC/0564).

# Supplementary materials

## Supplementary materials 1: EWAC coefficients (CSV file)

## Supplementary materials 2: Bayesian model report (PDF file)

## Supplementary materials 3: Subgroup analyses

Table 8: 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 9: 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\*\*\*

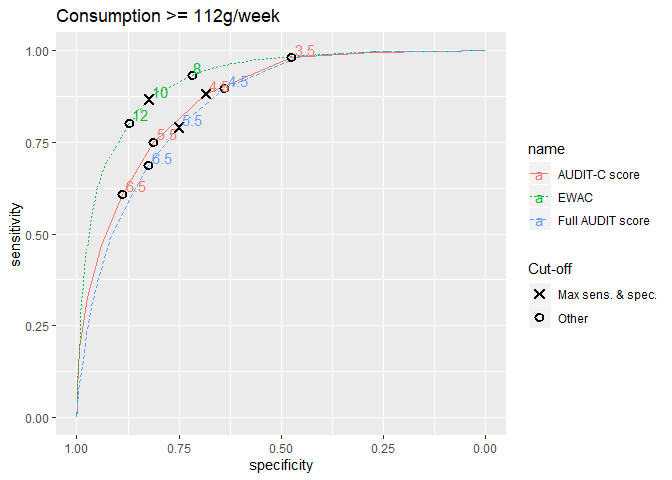
Notes:

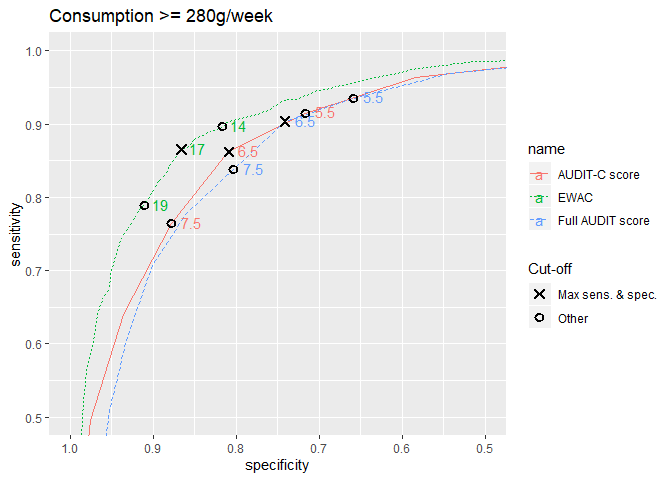
\*P < .05

\*\*P < .01

\*\*\*P < .001

## Supplementary materials 4: ROC curves





# Tables

Table 1: Overview of alcohol schedule used in this paper.

Survey module

Schedule

ATS QF module

On how many days, if any, did you personally drink a drink containing alcohol in the last four weeks?

What was the maximum number of units you personally consumed on any one day when drinking an alcoholic drink or drinks in the last four weeks?

On how many days, if any, in the last four weeks did you personally drink… [prompting in turn ‘51-60 units?’, ‘41-50 units?’, …, '1-2 units?

Health Survey for England

Thinking now about all kinds of drinks, how often have you had an alcoholic drink of any kind during the last 12 months? [8 items from ‘Almost every day’ to ‘Not at all in the last 12 months’]

Did you have an alcoholic drink in the seven days ending yesterday?

On how many days out of the last seven did you have an alcoholic drink?

Which day last week did you last have an alcoholic drink/have the most to drink?

Thinking about last [answer to previous question], what types of drink did you have that day? [list of 8 types of alcohol beverages]

[running through each type of beverage and recording number of units drunk]

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

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

[4] 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).

[5] 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).

[6] Australian Government. National Health and Medical Research Council. Australian Guidelines to Reduce Health Risks from Drinking Alcohol. 2009.

[7] UK Department of Health. UK Chief Medical Officers’ Low Risk Drinking Guidelines. August 2016. 2016.

[8] Wood AM, Kaptoge S, Butterworth AS, Willeit P, Warnakula S, Bolton T, et al. Risk thresholds for alcohol consumption: combined analysis of individual-participant data for 599912 current drinkers in 83 prospective studies. The Lancet 2018;391:1513–23. doi:[10.1016/S0140-6736(18)30134-X](https://doi.org/10.1016/S0140-6736(18)30134-X).

[9] 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.

[10] 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).

[11] 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).

[12] 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).

[13] 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).

[14] 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).

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

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

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

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

[19] 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).

[20] 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).

[21] 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).

[22] 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).

[23] 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).

[24] 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).

[25] Grant KA, Tonigan JS, Miller WR. Comparison of three alcohol consumption measures: A concurrent validity study. Journal of Studies on Alcohol 1995;56:168–72. doi:[10.15288/jsa.1995.56.168](https://doi.org/10.15288/jsa.1995.56.168).

[26] Hoeppner BB, Stout RL, Jackson KM, Barnett NP. How good is fine-grained Timeline Follow-back data? Comparing 30-day TLFB and repeated 7-day TLFB alcohol consumption reports on the person and daily level. Addictive Behaviors 2010;35:1138–43. doi:[10.1016/j.addbeh.2010.08.013](https://doi.org/10.1016/j.addbeh.2010.08.013).

[27] Vinson DC, Reidinger C, Wilcosky T. Factors affecting the validity of a Timeline Follow-Back interview. Journal of Studies on Alcohol 2003;64:733–40. doi:[10.15288/jsa.2003.64.733](https://doi.org/10.15288/jsa.2003.64.733).

[28] 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.

[29] 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).

[30] 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).

[31] 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).

[32] 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).

[33] 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).

[34] 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).

[35] 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).

[36] 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).

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

[38] 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).

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