medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247106; this version posted December 11, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

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

Authors

Dutey-Magni, PF,[1,4,*] Brown, J,[2] Holmes, J,[3] Sinclair, JMA[4]

1 Institute of Health Informatics, University College London, London, UK
2 Institute of Epidemiology and Health Care, University College London, London, UK
3 School of Health and Related Research, University of Sheffield, Sheffield, UK
4 Faculty of Medicine, University of Southampton, London, UK

*** Corresponding author:**

Peter F. Dutey-Magni
Institute of Health Informatics
University College London
222 Euston Road
London
NW1 2DA
UK
Telephone: (+44) 020 3108 6424
Email: p.dutey-magni@ucl.ac.uk

Running head

Estimator of Weekly Alcohol Consumption EWAC

Word count

3496 words

Protocol registration

Dutey-Magni, PF, Sinclair, JMA, Brown, J. 2018. 'Concurrent Validity of an Estimator of Weekly Alcohol Consumption (EWAC) Based on the Extended AUDIT.' *OSF*. November 12. doi:10.17605/OSF.IO/7WE4M.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247106; this version posted December 11, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

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: Cross-sectional diagnostic study based on 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 characteristic 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

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247106; this version posted December 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

1 Introduction

Alcohol consumption is responsible for 5% of disability-adjusted life years [1]. This burden extends far beyond the health burden of *alcohol use disorders*, as defined in the International Statistical Classification of Diseases (ICD-10 F10.1/F10.2 [2]) or the Diagnostic and Statistical Manual of Mental Disorders [3]. Clinical guidelines aiming to prevent [4], treat and reduce [5] harm from alcohol consumption recommend systematic screening for alcohol consumption using a validated diagnostic tool. However, conceptual differences (exemplified by the diagnostic classifications above) remain in how best to diagnose, measure, and communicate harm [6].

A global standard has emerged in the 10-item Alcohol Use Disorders Identification Test (AUDIT) [7]. Its first three questions, known as AUDIT-C and focussing on consumption, have equivalent predictive capability [8]. AUDIT-C is easy to use for patients and clinicians alike, making it an attractive choice for alcohol interventions in healthcare and other settings. Yet, AUDIT-C exhibits two limitations:

1. **Ceiling effect:** AUDIT-C's maximum response options for alcohol consumption frequency and quantity are heavily right-censored (Table 1). This creates a ceiling effect making the AUDIT-C poorly responsive to change in individuals with a high baseline score, even if they reduce consumption by up to thirty percent (e.g. frequency of drinking down from 7 to 5 days or quantity down from 16 to 11 drinks per day).

2. **AUDIT score intelligibility:** when presented alone, the AUDIT-C's ordinal score (range: 0–12) is not easily interpretable to patients and clinicians alike, who must be refer to cut-offs informed by diagnostic accuracy studies [9] and set with great variation internationally [10]. The fact that the AUDIT-C score is not easy to interpret poses challenges in the delivery of brief interventions. Healthcare professionals lack confidence in discussing alcohol consumption [11–13], and the abstract AUDIT-C score requires training on how to deliver feedback [11]. Similar challenges are faced when

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247106; this version posted December 11, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

communicating AUDIT-C results in self-administered interventions such as web apps [14].

The 'Extended AUDIT-C' addresses the first of these limitations thanks to a greater range of response options on quantity and frequency (Table 1). It has been used in UK research as part of two trials [15,16] and one continuous household survey [17] to measure characteristics of consumption that could not have been measured with the right-censored AUDIT-C.

The present study proposes to address the second of these limitations, by developing and validating a fast Estimator of Weekly Alcohol Consumption (EWAC) computed from the Extended AUDIT-C. The EWAC is designed to offer equivalent diagnostic capabilities in relation to alcohol use disorders (being based on the AUDIT-C), while providing a continuous and more directly interpretable metric of alcohol consumption. This is intended facilitate the delivery of brief interventions, since measuring alcohol consumption is a crucial part of behaviour change techniques (self-monitoring, feedback on behaviour, social comparison) commonly employed in self- [18] and clinician-administered [19] interventions.

This structure of the paper is as follows. The paper reports the development of the EWAC, which is the product of quantity (AUDIT-1) by frequency (AUDIT-2) with adjustment for occasional intense drinking (AUDIT-3). Coefficients are estimated in a hierarchical Bayesian response model using Extended AUDIT-C and Graduated-Frequency (GF) data from a large English household survey, the Alcohol Toolkit Study [17]. The same data are then used to compare EWAC and GF, derive metrics of bias and error, and detect potential subgroup variations in accuracy. Finally, the paper tests the equivalence of receiver operating characteristics of EWAC compared with AUDIT-C and the full AUDIT in predicting alcohol consumption exceeding 112g/week (14 UK units) and 280g/week (35 UK units).

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247106; this version posted December 11, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

2 Methods

2.1 Design

Neither the AUDIT-C nor the Extended AUDIT-C provide a direct measure of alcohol consumption. Instead, we employ methods developed for quantity-frequency-variability instruments [20]. For every individual i, the EWAC is computed as the product of F_i and Q_i (AUDIT questions 1 and 2 respectively) adjusted with the frequency of intense drinking V_i (AUDIT-3):

$$\mathrm{EWAC}_i = F_i Q_i + V_i b$$

where b denotes the average units of alcohol consumed in an intense drinking day.

Coefficients F, Q, V and b are unknown. The present study considers two sets of candidate coefficients:

- AUDIT response item interval midpoint (e.g. 2.5 for '2 to 3 times per week')
- 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 $\mathrm{GF}_i = F_i Q_i + V_i b + e_i$, where e denotes independently normally distributed errors. We set parabola-shaped informative priors on coefficients F, Q, V. Details on model fitting, convergence evaluation and prior tuning is reported in supplementary information S1.

2.2 Participants

Data originate from baseline measures in waves 110–133 (November 2015–October 2017) of the Alcohol Toolkit Study (ATS), a computer-assisted personal interview of residents of private English households aged ≥16 years [17]. Participants were included in the analysis if they completed both the Extended AUDIT and the GF questionnaires. Out of 40,832 participants, 14,408 (35%) reported 'never' consuming alcohol in AUDIT question 1 and were not asked any further AUDIT or GF questions. They were thus excluded.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247106; this version posted December 11, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

2.3 Measures

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

The reference standard used is the ATS GF schedule (supplementary information S3), in which participants described how many times they consumed given quantities of alcohol *during the last 4 weeks* [21]. A longstanding obstacle in alcohol research and treatment lies in the absence of undisputed 'gold standard' or biomarker for objectively determining alcohol consumption. Instead, a number of instruments measure self-reported consumption with varying validity and reliability over different durations. Comprehensive reviews [22–27] indicate that yesterday recall and prospective diaries tend to record higher (and more accurate) alcohol consumption by minimising recall bias, followed by GF measures. The GF schedule's main advantage lies in measuring occasional heavy consumption, which can constitute an important proportion of total consumption. Although widely employed in population surveys, the GF schedule is uncommon in clinical practice.

We examined the EWAC's coverage of per-capita alcohol retail sales for England in 2014 [28], which capture all alcohol produced/processed in or imported to England for sale or consumption.

We also cross-examined the EWAC's distribution against other reference estimates obtained from 8,610 household residents aged ≥16 years participating in the 2011 Health Survey for England (HSE) [29]. On that particular year, the HSE included (a) a computer-assisted interviewer-led beverage-specific quantity-frequency questionnaire; (b) a yesterday recall question [27]; and (c) a prospective 7-day diary [30].

2.4 Missing data

We excluded 4,020/26,424 participants (15.2%) who did not have a valid GF alcohol consumption record (Table S7). In the remaining 22,404 valid observations, GF data were

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247106; this version posted December 11, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

assumed to be missing at random conditionally on the Extended AUDIT-C responses. In subgroup analyses, a further 530 observations (2.4%) assumed to be missing at random were excluded.

2.5 Analyses

The protocol was pre-registered [31]. Results are reported in UK alcohol units (8g or 10mL of pure alcohol). Analyses were conducted in R [32–34]. Scripts are available online [35].

2.5.1 Overall bias and error

The agreement between the EWAC and the GF was quantified as follows:

- bias was estimated by the **mean deviation** to the reference standard $MD = n^{-1} \sum_{i=1}^{n} (\mathrm{EWAC} - \mathrm{GF})$. We tested the hypothesis that the MD does not exceed 1 UK unit using a two-sided t-test.
- precision was estimated by the **root mean squared deviation** $RMSD = \sqrt{n^{-1} \sum_{i=1}^{n} (\mathrm{EWAC} - \mathrm{GF})^2}$, a measure of total error capturing 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 tested the hypothesis that the RMSD does not exceed 2 UK units using a one-sided χ^2 homogeneity test.

Two sets of candidate coefficients were considered (section 2.1). We only report findings for the candidate set producing the lowest bias and error.

2.5.2 Subgroup bias and error

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

- the simple deviation $(\mathrm{EWAC} - \mathrm{GF})$ was regressed in a linear model to test subgroup differences in MD

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247106; this version posted December 11, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

- the squared deviation $(\mathrm{EWAC} - \mathrm{GF})^2$ 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; these are interpreted as the ratio of the subgroup RMSD to the reference RMSD, a ratio >1 indicating worse precision than in the reference category.

Both models (supplementary Tables S5–S6) included the following predictors: sex by age group; ethnic group; highest educational qualification; religion; smoking status. Additional models (supplementary Table S6) were fitted solely in respondents with an AUDIT-C score ≥5 or an AUDIT score ≥8, 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.

2.5.3 Receiver Operating Characteristics

We tested the EWAC's superiority to the traditional AUDIT and AUDIT-C scores in predicting consumption exceeding 14 or 35 UK units/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≥8) which is above 35 units for women and 50 units for men [36]. We tested the hypothesis that the EWAC has an identical receiver operating characteristic full area under the curve (AUC) to the AUDIT-C and the full AUDIT scores using nonparametric paired AUC tests [37]. AUDIT-C and AUDIT scores were calculated from the Extended AUDIT by capping the contribution of each question to 4.

2.5.4 Aggregate concurrent validity

We compared the empirical cumulative distributions of (1) the EWAC computed in the ATS; (2) the GF estimator in the ATS; (3) the beverage-specific estimator in the 2011 HSE; (4) the prospective diary estimator in the 2011 HSE in adults aged ≥18 years. We report the proportions of on-trade and off-trade alcohol sales [28] accounted for by each method. Poststratification survey weights adjusted for nonresponse bias in sources (1-3), and self-selection into prospective diary data collection in source (4).

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247106; this version posted December 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

3 Results

3.1 Overall bias and accuracy

EWAC coefficients estimated empirically (supplementary information S4) had smaller bias and error and were used for the remainder of the analysis. With those, the EWAC's Pearson's correlation with GF was estimated at r = 0.71 [0.71, 0.72] (Kendall's rank correlation τ = 0.63).

The mean deviation (MD) was 0.2 alcohol units/week [95% CI: 0.04, 0.3]. This bias is smaller than the preregistered $\pm$1-unit bias tolerance (p = 1.000).

The root mean squared deviation (RMSD) at 10.9 units/week [95% CI: 9.8, 12.0] was significantly greater than the pre-registered 2-unit total error tolerance (p<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). In 50% of participants, the EWAC fell within $\pm$2 UK units of the GF weekly consumption estimate. Thus, an interval defined as the EWAC $\pm$2 units (e.g. '10 to 14 units') would contain the reference standard for half of individuals, while an interval ranging $\pm$3 units (e.g. '9 to 15 units') would contain the reference standard for 60% of individuals.

Plots of EWAC against GF (Figure 1) indicate a slight positive bias for consumptions up to 10–14 units/week, and a slight negative bias beyond. The EWAC only starts losing 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)).

3.2 Subgroup accuracy

MDs and RMSDs were regressed against respondent characteristics to identify potential subgroup differences in bias or precision from the reference category (White British females aged 25–34 years without educational qualifications, who never smoked). Model predictors accounted for <2% of the variance, indicating stability in MD and RMSD across subgroups (Table S5).

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247106; this version posted December 11, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

Figure 2 shows subgroups with a predicted MD exceeding $\pm$1 unit/week and a p-value<0.05. Respondents of Black/Other/White Other ethnic groups had an overestimated EWAC: their MDs exceeded the reference MD by 2.0 units [0.9, 3.1]; 2.7 units [0.8, 4.5]; and 1.1 units [0.5, 1.7] respectively. MDs of men aged 55–64 years, or ≥75 years exceeded the reference MD by 1.3 units [0.2, 2.4] and 1.5 units [0.3, 2.8] respectively. There was no evidence of an effect of favourite drink or attempts to reduce alcohol intake in increasing-risk drinkers (Table S6).

Figure 3 shows subgroups with an RMSD at least 25% smaller/greater than the reference. RMSD was 58% [50, 67%] greater in current smokers and 34% [14, 56%] greater in respondents who stopped smoking in the past year. It was also 44% [29, 60%] and 34% [19, 50%] greater respectively in men and respondents aged 16–24 years. In contrast, error was 35 to 70% smaller in Black/Other/White Other ethnic groups.

Figure 4 presents the same metrics for increasing-/higher-risk drinkers (AUDIT-C≥5 or AUDIT≥8). Educational qualifications significantly improved the agreement between EWAC and the reference standard in this group (Table S7). School and degree-level qualifications reduced RMSD by 24% [12, 37%] and 37% [23, 52%] respectively, suggesting that respondents may better evaluate and recall alcohol beverage content. Conversely, the RMSD of respondents who attempted to reduce alcohol use was 23% [16, 30%] larger than the reference RMSD.

3.3 Receiver Operating Characteristics

We examined the EWAC's ability to predict consumption exceeding 14 or 35 UK units/week. The full areas under the receiver operating characteristic curves (AUC, Figure S7) are presented along sensitivity and specificity at the best thresholds in Tables 3–4.

14 UK units/week (increasing-risk): EWAC increases the AUC by 5 percentage points compared to the AUDIT-C score (p<0.001); and 7 percentage points compared to the full AUDIT score (p<0.001). The cut-off maximising the sum of specificity and sensitivity on the

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247106; this version posted December 11, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

EWAC is 10 units/week. The sensitivity at this threshold is identical to AUDIT-C, but specificity gains 13 percentage points. Using the nominal cut-off of 14 units/week on the EWAC raises specificity to 0.922, at the cost of a reduction in sensitivity to 0.705 (Table 3).

35 units/week (higher-risk): EWAC provides small increases in AUC compared with the AUDIT-C score (p<0.001) and the full AUDIT score (p<0.001). The best cut-off for detecting consumption in excess of 35 units/week using the EWAC is 17 units/week (Table 4).

3.4 Empirical distribution

Table 5 estimates adult residents' total alcohol consumption in England using four different sources, and compares them with alcohol retail sales. The HSE exhibits the highest estimates and coverage of alcohol sales. EWAC amounts to 71% of the total consumption estimated by the HSE's prospective diary, and 48% of retail sales.

Figure 5 suggests that the EWAC, like the ATS GF, estimates a greater prevalence of lower-risk (≤14 units/week) and increasing-risk alcohol use than HSE.

4 Discussion

4.1 Main findings

We developed a continuous metric of estimated weekly alcohol consumption using the 3-question Extended AUDIT-C. In a representative household survey, this metric estimated usual alcohol consumption with a mean precision of ±11 units/week and a median precision of ±2 units when compared with the GF reference standard.

4.2 Strengths and limitations

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 study [38] previously reported mean consumption by AUDIT-C score, but without quantifying bias or precision of such a measure. Other studies have evaluated the AUDIT-C's accuracy in estimating alcohol consumption, but only in relation to predicting consumption in excess of predefined thresholds [9]. Such

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247106; this version posted December 11, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

studies achieved AUCs ranging 0.83–0.96. Our study found the EWAC to be superior to both the AUDIT-C and the full AUDIT in predicting GF exceeding 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 and extensive range of subgroup analyses reported. Bias was mostly consistent across subgroups examined (age/sex, education, smoking status, religion), with one exception. EWAC overestimated alcohol consumption by 2-3 UK units/week in Black/Other ethnic groups; though overall precision was better in Black/Asian/Other ethnic groups than in White British respondents. Variation in the sensitivity of AUDIT-C across ethnic groups has previously been noted in the US [39].

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 [7]. 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, based on the last 30 years of international research. For instance, a previous study which found the AUDIT-C to be responsive to changes of 70g/week [40] can suggest that the EWAC's own responsiveness to change should be equivalent, if not greater than the AUDIT-C's, given the Extended AUDIT-C's additional response items.

Despite this, the EWAC's design does not escape limitations shared by all methods of screening or categorising alcohol use disorders. The conceptualisation of alcohol use disorders is related to, but does not exclusively depend upon the amount of alcohol consumed. Since Jellinek's description of 'the disease concept of alcoholism' [41] there have been numerous attempts to categorise the range of phenotypes characterising alcohol use disorders in the absence of any biomarker to 'verify' the presence of a particular pathology. The EWAC, by limiting itself to an estimation of alcohol consumption is transparent across a

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247106; this version posted December 11, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

wide range of alcohol use disorders but does not measure the other factors underpinning this complex and heterogeneous condition [6,42].

We note other study limitations. Our findings may not apply to other countries, or to subpopulations with an atypical alcohol consumption (e.g. patients seeking care for severe alcohol dependence, alcohol-related liver disease). Like the HSE, the ATS does not cover populations excluded from most sampling frames (e.g. communal/carceral institutions, homeless/no fixed abode populations).

We also note that the reference standard, as all self-reported measures, is imperfect. This may introduce bias into estimates of precision: by definition, the reference standard's own error will inflate the RMSD. In other words, it is likely that a proportion of the RMSD is attributable to error in the GF measures rather than the EWAC.

Previous research offers reassurance that the EWAC's error is comparable to that of other (even the more time-consuming) questionnaires. One telephone interview study [43] measured a Pearson's correlation r = 0.86 and 0.87 in drinking frequencies and volumes between a 12-month GF interview and a subsequent 28-day prospective diary. A postal study [44] measured a correlation r = 0.89 between a 30-day GF and a prospective diary (identical reference periods). Heeb *et al.* [22] measured a Kendall's rank correlation τ = 0.41 between a 7-day GF and a 7-day prospective diary. For reference, our study measured a correlation coefficient r = 0.71 and τ = 0.63 between a 6-month EWAC and a 4-week GF and suggests that EWAC's agreement with GF is similar to that of diaries.

4.3 Potential applications

Being equivalent or superior to the AUDIT in speed, accuracy, and international standardisation, the EWAC appears suitable for use in clinical practice to support brief interventions and 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

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247106; this version posted December 11, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

EWAC is already compatible with medical records information models developed in the Systematised Nomenclature of Medicine Clinical Terms (SNOMED CT, Alcohol intake (observable entity) [45]) and by the English Royal College of Physicians [46]. Such information can have secondary uses as a variable in other disease risk scores, or to prospectively recording of long-term alcohol exposure, an important risk factor for a range of medical conditions.

Compared to an AUDIT-C score, alcohol consumption is a relatable scale, with health education/promotion qualities and potential to support interventions focusing on recognising the alcohol content and volume of drinks. Knowledge of alcohol beverage content is generally poor [47], and many countries have not adopted a measure of standard drinks [48]. Two thirds of English drinkers could assess the standard drink equivalent of one alcohol unit [49].

Assessment of alcohol consumption is not well embedded in clinical practice [50]. The EWAC calculator fills a gap in resources empowering individuals to understand, monitor and control their alcohol consumption with–or without–the involvement of healthcare professionals. The EWAC appears particularly suitable for digital interventions.

Nutt, Rehm *et al.* [51,52] argued 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 EWAC's dimensional rather than categorical format may facilitate this while avoiding the stigma sometimes associated with clinical categorisations of alcohol use disorders.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247106; this version posted December 11, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

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 EWAC online calculator development was funded by the Wessex Academic Health Science Network. 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. PD, JH and JMAS declare no competing interests.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247106; this version posted December 11, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

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.

[2] World Health Organisation. ICD-10: International statistical classification of diseases and related health problems: Tenth revision. 2004.

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

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

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

[6] 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.

[7] 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.

[8] 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.

[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.

[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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] 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.

[18] 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.

[19] 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.

[20] 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.

[21] 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.

[22] 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.

[23] 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.

[24] 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.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247106; this version posted December 11, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

[25] 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.

[26] 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.

[27] 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.

[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] 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.

[30] 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.

[31] 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.

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

[33] Wickham H. Tidyverse: Easily install and load the 'tidyverse' 2017.

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

[35] Dutey-Magni PF. Concurrent validity of an Estimator of Weekly Alcohol Consumption (EWAC) based on the Extended AUDIT: Computer scripts 2020. doi:10.5281/zenodo.4315024.

[36] 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.

[37] 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.

[38] 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.

[39] 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.

[40] 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.

[41] Jellinek EM. The disease concept of alcoholism. New Haven: Hillhouse Press; 1960. doi:10.1037/14090-000.

[42] 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.

[43] 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.

[44] 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.

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

[46] 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.

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

[48] 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.

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

[50] 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.

[51] 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.

[52] 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.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247106; this version posted December 11, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

Figures and tables

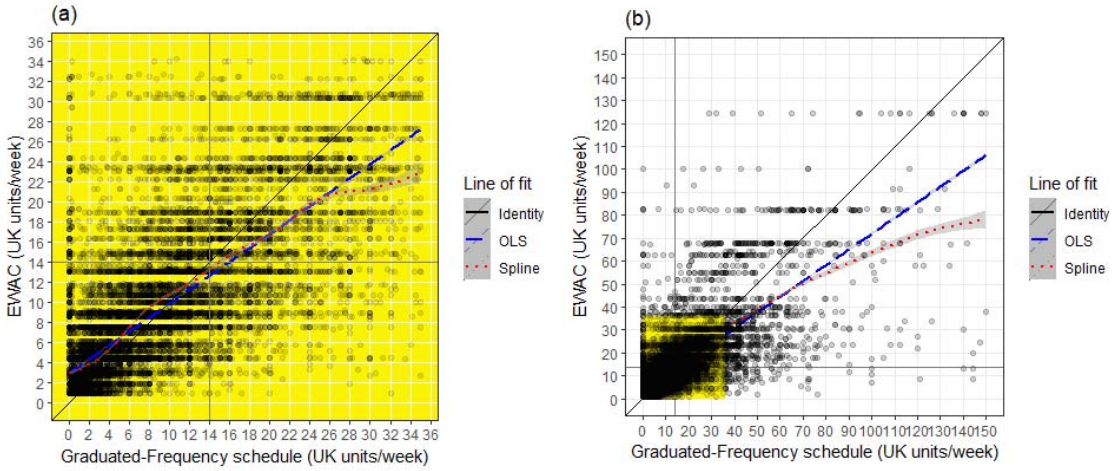

Figure 1. Plots of EWAC against GF in (a) low/increasing risk ATS respondents (n=21,338) and (b) all ATS respondents (n=22,373)

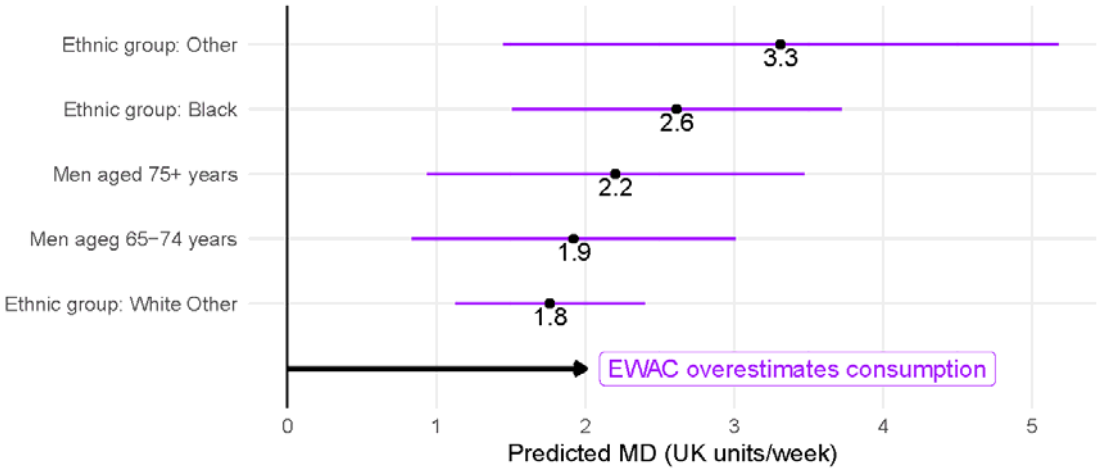

Figure 2. Forest plot of modelled MD for selected subgroups

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247106; this version posted December 11, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

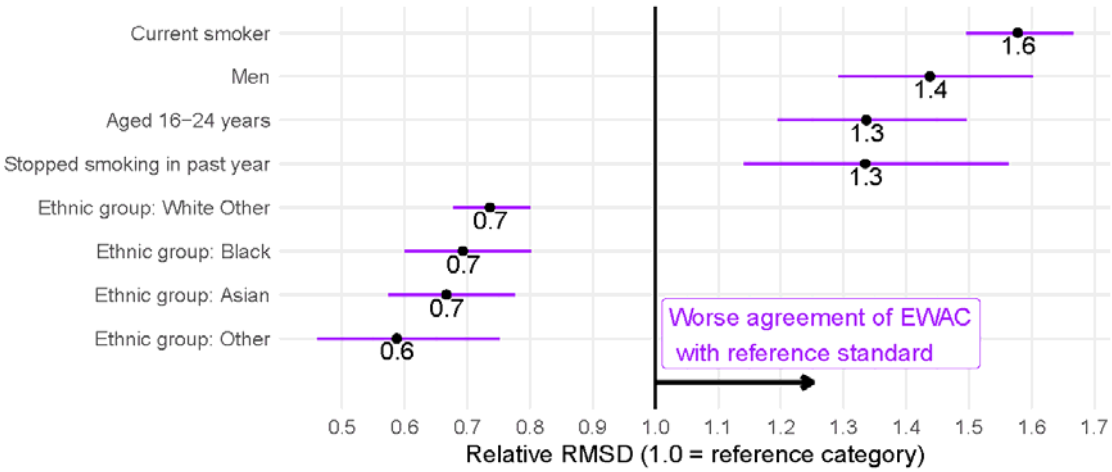

Figure 3. Forest plot of RMSD ratio (selected subgroups to reference category)

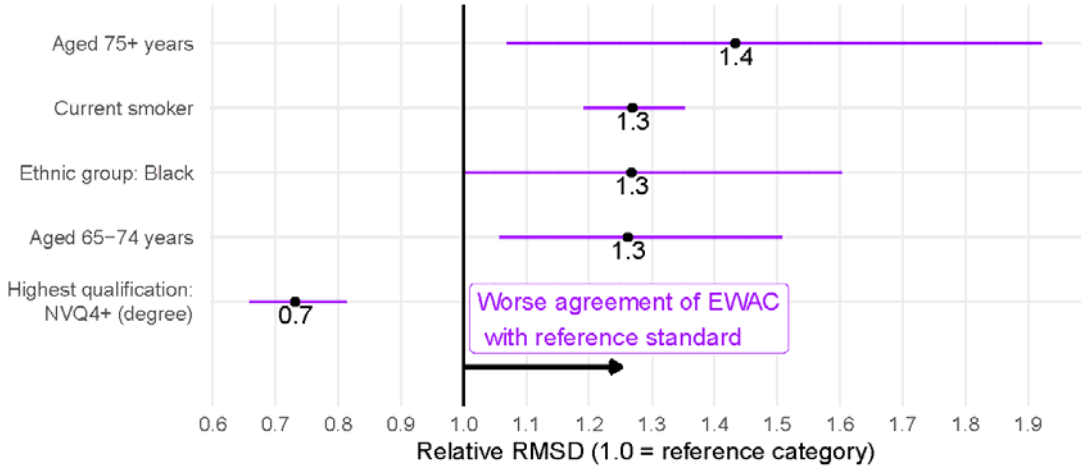

Figure 4. Forest plot of RMSD ratio (selected subgroups to reference category) in respondents with a hazardous/harmful alcohol use (AUDIT-C>=5 or AUDIT>=8)

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247106; this version posted December 11, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

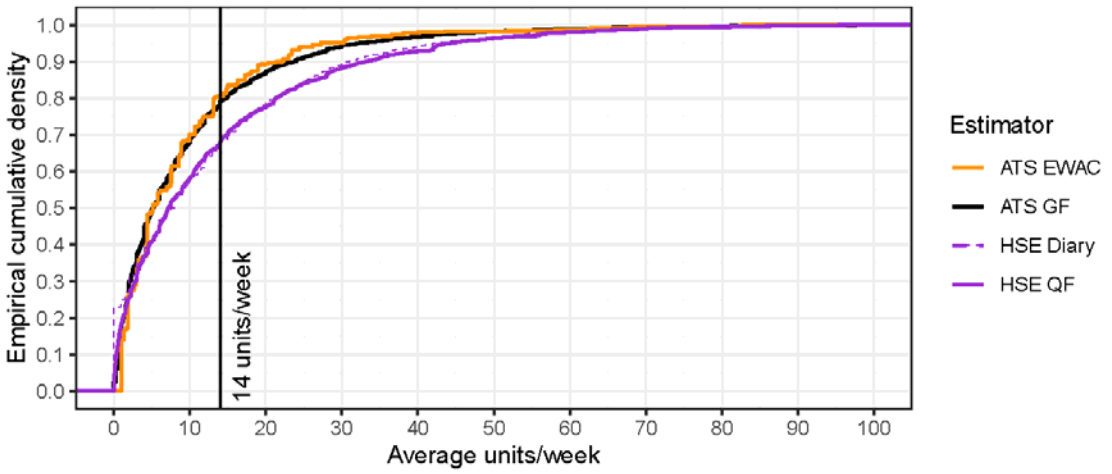

Figure 5. Empirical cumulative distribution function of weekly alcohol consumption in England according to four alcohol schedules in residents aged 18 years and over

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

	Response items						
AUDIT-1: 'How often do you have a drink containing alcohol?'							
AUDIT-C	Never	Monthly or less	2 - 4 times per month	2 - 3 times per week	4+ times per week	–	
Extended AUDIT-C	Never	Monthly or less	2 - 4 times per month	2 - 3 times per week	4-5 times per week	6+ times per week	
AUDIT-2: 'How many units of alcohol do you drink on a typical day when you are drinking?'							
AUDIT-C	0 - 2	3 - 4	5 - 6	7 - 9	10+	–	–
Extended AUDIT-C	0 - 2	3 - 4	5 - 6	7 - 9	10-12	13-15	16+

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

Absolute deviation percentiles	**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

medRxiv preprint doi: https://doi.org/10.1101/2020.12.11.20247106; this version posted December 11, 2020. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

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, 0.860]	5.5	0.791	0.751
EWAC	0.921	[0.917, 0.925]	10	0.876	0.816

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

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.

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

Study	**Mean** (UK units/week)	**Median** (UK units/week)	**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	–	–	–	–