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

* [1 Introduction](#introduction)
* [2 Methods](#methods)
  + [2.1 Approach](#approach)
  + [2.2 Data sources](#data-sources)
  + [2.3 Study 1: EWAC estimation and concurrent validity](#study-1-ewac-estimation-and-concurrent-)
  + [2.4 Study 2: Concurrent validity in hospital outpatients](#study-2-concurrent-validity-in-hospital)
  + [2.5 Study 3: Aggregate concurrent validity](#study-3-aggregate-concurrent-validity)
* [3 Results](#results)
  + [3.1 Study 1: EWAC estimation and validation](#study-1-ewac-estimation-and-validation)
    - [3.1.1 Overall bias and accuracy](#overall-bias-and-accuracy)
    - [3.1.2 Subgroup accuracy](#subgroup-accuracy)
    - [3.1.3 Receiver Operating Characteristics](#receiver-operating-characteristics)
  + [3.2 Study 2: Concurrent validity in hospital patients](#study-2-concurrent-validity-in-hospital)
  + [3.3 Study 3: Empirical distribution functions](#study-3-empirical-distribution-function)
* [4 Discussion](#discussion)
  + [4.1 Main findings](#main-findings)
  + [4.2 Strengths and limitations](#strengths-and-limitations)
  + [4.3 Limitations](#limitations)
  + [4.4 Is the EWAC good enough to be used?](#is-the-ewac-good-enough-to-be-used)
  + [4.5 What are the potential applications?](#what-are-the-potential-applications)
* [5 Conclusion](#conclusion)
* [6 Ethics](#ethics)
* [7 Supplementary materials](#supplementary-materials)
  + [7.1 Supplementary materials 1: EWAC coefficients (CSV file)](#supplementary-materials-1-ewac-coeffici)
  + [7.2 Supplementary materials 2: Bayesian model report (PDF file)](#supplementary-materials-2-bayesian-mode)
  + [7.3 Supplementary materials 3: Subgroup analyses](#supplementary-materials-3-subgroup-anal)
  + [7.4 Supplementary materials 3: ROC curves](#supplementary-materials-3-roc-curves)
* [8 Appendix hospital study](#appendix-hospital-study)
* [9 Tables](#tables)
* [References](#references)

# 1 Introduction

Globally, alcohol consumption is responsible for 5% of all deaths and disease disease [[1](#ref-Shield2020)]. The burden of alcohol consumption goes far beyond the burden of alcohol use disorders, defined in the International Statistical Classification of Diseases (ICD-10 F10.1/F10.2; ICD-11 QE10/6C40.1).

A number of strategies exist in order to prevent [[2](#ref-NICE-PH24)], or treat and reduce [[3](#ref-NICE-CG115)] harm from alcohol consumption. Some recommend systematic screening of alcohol use disorders using a validated diagnostic tool. More recently, others have advocated monitoring alcohol consumption like many other risk factors (eg body mass index or cholesterol), making individuals more conscious of their current intake and associated risk [[4](#ref-Nutt2014)]. This approach is intended to promote timely behaviour change interventions (and, if necessary, pharmacological interventions) and prevent the emergence of alcohol use disorders. Recent public health recommendations to the public [[6](#ref-AU.alcoguidelines2009)] also now promote the idea of reducing consumption in the absence of a ‘safe’ level of alcohol consumption.

In contrast with alcohol use disorder screening tools, alcohol consumption assessment tools are not well embedded in clinical practice. Potential reasons include: lack of diagnostic validation, difficulty to self-assess alcohol consumption, and time pressures. This creates a gap in the range of resources destined to help individuals understand, monitor and take control of their alcohol consumption with–or without–the involvement of primary care physicians.

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) [[8](#ref-Babor2001)], a 10-item questionnaire schedule widely employed in clinical practice and clinical research as a diagnostic test for alcohol use disorder. The short, 3-item consumption version known as AUDIT-C has good diagnostics accuracy for both: (a) interview-based clinical diagnoses; and (b) consumption in excess of maximum recommended intakes (e.g. 140g/week in Australia, 112 g/week in UK) in a variety of populations [[9](#ref-DeMeneses-Gaya2009)]. The present study aims to validate the AUDIT-C 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, thanks to a greater range of response options on quantity and frequency (items 1 and 2). The Extended AUDIT-C has been used in UK research as part of two trials [[10](#ref-Kaner2013c)] and one continuous household survey [[12](#ref-Beard2015a)] to capture greater information on the higher risk drinkers, based on the observation that AUDIT consumption items are right-truncated.

This paper reports three distinct empirical studies. 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 across a number of demographic subpopulations in England. Study 2 tests the concurrent validity of the EWAC compared with the 28-day Timeline Followback in a clinical population (visitors of an acute hospital). Study 3 compares the population-wide total and empirical cumulative distribution of alcohol consumption in England using the EWAC, the ATS graduated frequency schedule, the Health Survey for England, and official statistics on alcohol sales.

# 2 Methods

## 2.1 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 measure of usual alcohol consumption, but the product of frequency of drinking (AUDIT-C item 1) and quantity of drinking (AUDIT-C item 2) can be used to estimate usual alcohol consumption, with adjustment for occasional heavy use (AUDIT-C item 3), following methods developed for quantity-frequency-variability instruments [[13](#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 \(b\) 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
* using a statistical model to estimate the coefficients directly from gold standard data.

All results are reported in UK alcohol units (8g or 10mL of pure alcohol). Analyses are conducted in R [[14](#ref-RCoreTeam2017)] using packages tidyverse, survey, rstan [[15](#ref-package-tidyverse)–[18](#ref-Lumley2004)]. Computer scripts for all analyses are available on an online repository [[19](#ref-Dutey2020)].

## 2.2 Data sources

A longstanding obstacle in alcohol research and care is the absence of a diagnostic gold standard, which is not addressed even by the development of new biomarkers. Instead, a number of instruments exist which measure self-reported alcohol consumption with varying validity and reliability. A state-of-the art review from [[20](#ref-Greenfield2000)] is summarised in Table 1 below.

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 |

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 [[21](#ref-Heeb2005)].

The present paper employs four sources of data. An overview is available in Table 2 below, together with references to publicly available methodological descriptions.

Table 2: Description of data sources

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Source | Time | Study population | Measurements | Validation | Sample size | Design |
| Alcohol Toolkit Study (ATS) [[12](#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 | – | 40,832 | TBC |
| Health Survey for England (HSE) [[23](#ref-NatCenSocialResearch2013)] | 2011 | Residents of private English households aged 18 years and over | Interviewer-led beverage-specific QF | This UK-specific schedule is consistent with a prospective diary [[24](#ref-Boniface2014)] and yesterday recall [[25](#ref-Stockwell2016)]. | 8,610 | Multistage random sampling |
| 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 [[26](#ref-Grant1995)]. This recall bias increases with the number of days elapsed [[27](#ref-Hoeppner2010),[28](#ref-Vinson2003)] | 130 | Cross-sectional recruitment with block randomisation to (a) self-administered AUDIT-C or (b) researcher-administered AUDIT-C. |
| Alcohol retail sales [[29](#ref-PHE2017)] | 2014 | English population aged 18 years and over | – | [[29](#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 |

## 2.3 Study 1: EWAC estimation and concurrent validity

Study 1 evaluates the accuracy of the EWAC derived from a researcher-administered Extended AUDIT-C in general household population. The gold standard is the GF schedule from the ATS data source. We compare two sets of coefficients: AUDIT response item interval midpoint; and coefficients estimated using a statistical model (detail reported in supplementary material).

First, the discrepancy between the EWAC and the GF schedule is quantified, by treating GF as a gold standard. It enables the examination of two types of deviations across every participant \(i\) in the ATS sample:

* the mean deviation \(\text{MD} = n^{-1} \sum\_{i=1}^{n}{(\rm{EWAC}\_i - \rm{GF}\_i )}\) can be assimilated to a measure of bias under the assumption that GF is a gold standard. In this study, we test whether the MD is greater than 1 UK unit using a two-sided \(t\)-test.
* the root mean squared deviation \(\text{RMSD} = \sqrt{n^{-1} \sum\_{i=1}^{n}{( \rm{EWAC}\_i - \rm{GF}\_i )^2}}\) can be assimilated to a measure of accuracy or total error: it capture both bias and random deviation from the assumed gold standard. For example, an RMSD of 2 signifies that the EWAC is on average with +/-2 UK units of the gold standard. In this study, we test the hypothesis that the RMSD is greater than 2 UK units using a one-sided \(\chi^2\) homogeneity test.

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

* the simple deviation \((\text{EWAC}-\text{GF})\) 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 MSD. Model coefficients are transformed (exponential then square root transformations) in order to obtain the relative RMSD, that is, the ratio of RMSDs of the subgroup to the reference category. Both models include the following predictors: sex by age group; ethnic group; highest educational qualification; religion; smoking status. Additional models are estimated for 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), who provided additional information during the interview: favourite drink (beer; wine; spirits alone; mixed spirits; cider; other); and whether the respondent attempt 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 has equivalent or better predictive value in predicting drinking in excess of 14 UK units per week compared to the traditional AUDIT-C score. In the UK, a score of 5 or more is categorised as ‘increasing risk’ or ‘higher risk’, and regarded as indicating consumption in excess of 14 units per week. The study tests the hypothesis that the EWAC’s receiver operating characteristics’ area under the curve (AUC) is greater that the AUDIT-C score using a nonparametric paired AUC test [[30](#ref-Delong1988)].

The ATS data (\(n=\) 40,832) is affected by missing data: 35% of respondents (\(n=\) 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, a further 0.5% (530) repondents were assumed to have data missing at random and were excluded.

A pre-registered protocol for this study can be found in [[31](#ref-Dutey2018)].

## 2.4 Study 2: Concurrent validity in hospital outpatients

A second study aims to confirm the robustness of findings in (a) a patient population; (b) when the Extended AUDIT-C is self-administered; (c) using a different gold standard: the 28-day TLFB.

Participants were recruited from a range of clinics at a large acute hospital in Southampton, UK: orthopaedics outpatient, endoscopy day cases, young adult outpatient, infusion ???. Participants were randomised to one of two groups:

1. self-administered Extended AUDIT-C
2. researcher-administered Extended AUDIT-C using block randomisation.

A total of 130 participants were recruited in hospital clinics, with data successfully collected on the Extended AUDIT-C for 81% of participants (\(n\) = 105). The Extended AUDIT-C was administered first, with 44 participants randomised to self-administration using a pen and paper form, while the remaining 61 were administered by a researcher.

Once this first questionnaire completed, both groups were asked to complete a 28-day alcohol TFLB questionnaire administered by the researcher.

To model the relationship between the EWAC and the TLFB, the number of units consumed on any day was assumed to follow a negative binomial distribution, the rate of which is determined by the latent usual alcohol consumption as well as the following variables: AUDIT mode of administration (self-administered; researcher-administered), week day, number of days elapsed since (1-7; 8-14; 15-21; 22-28). TLFB daily number of units consumed were rounded to the nearest integer and regressed in a negative binomial regression model against the aforementioned predictor, as well as the EWAC divided by 7 to obtain its daily equivalent. The corresponding regression coefficient measures the ratio of TLFB to EWAC.

## 2.5 Study 3: Aggregate concurrent validity

We examine the aggregate consistency in alcohol consumption estimates across all residents of England aged 18 years and over by plotting the empirical cumulative distribution 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 MYPES; in (4), similar postratification weights adjusted for self-selection into participation to the prospective diary data collection. Finally, we examine the percentage of total alcohol sales for England accounted for by each method. Alcohol sales (both on-trade and off-trade) estimates for England (population aged 18 years and over) in 2014 were obtained from Public Health England [[29](#ref-PHE2017)].

# 3 Results

## 3.1 Study 1: EWAC estimation and validation

### 3.1.1 Overall bias and accuracy

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

The EWAC computed with the midpoint of the AUDIT item intervals produces a mean deviation (MD) of 0.71 UK alcohol units/week, indicating a bias inferior to the preregistered +/- 1-unit bias allowance (\(p\) = 1.000). The root mean squared deviation (RMSD) estimate of 12.3 units [95% CI: 11.2;13.2] is significantly greater than the pre-registered 2-unit total error allowance (\(p\) < 0.001), suggesting that the EWAC falls on average 12 units away from the GF gold standard.

Coefficients estimated empirically (statistical model reported in supplementary materials 2) provide a small improvement: MD = 0.18 (\(p\) = 1.000). Error remains statistically significantly greater than 2 (\(p\) < 0.001) with RMSD = 10.9 [95% CI: 9.8;12].

The RMSD masks a dispersed and skewed error pattern. Table 3 shows that, for 50% of participants, the EWAC falls within +/- 2 UK units of the GF weekly consumption estimate. In other terms, an interval estimate defined as the EWAC +/- 2 units (eg ‘10 to 14 units’) would be accurate for half of individuals, while an interval estimate defined as the EWAC +/- 3 units (eg ‘9 to 15 units’) would be accurate for 60% of individuals.

Table 3: Percentiles of the absolute deviation between EWAC and GF schedule (n = 22404)

`10%` `20%` `30%` `40%` `50%` `60%` `70%` `80%` `90%` `95%` `99%`  
 0.4 0.7 1 1.5 2.1 3 4.2 6.2 10.6 17 38.7

Figure 1 compares individual EWAC and GF values. We note the departure of lines of best fit from the diagonal, demonstrating the EWAC’s small but consistent positive bias (MD > 0). This discrepancy is more severe beyond 25 units per week.

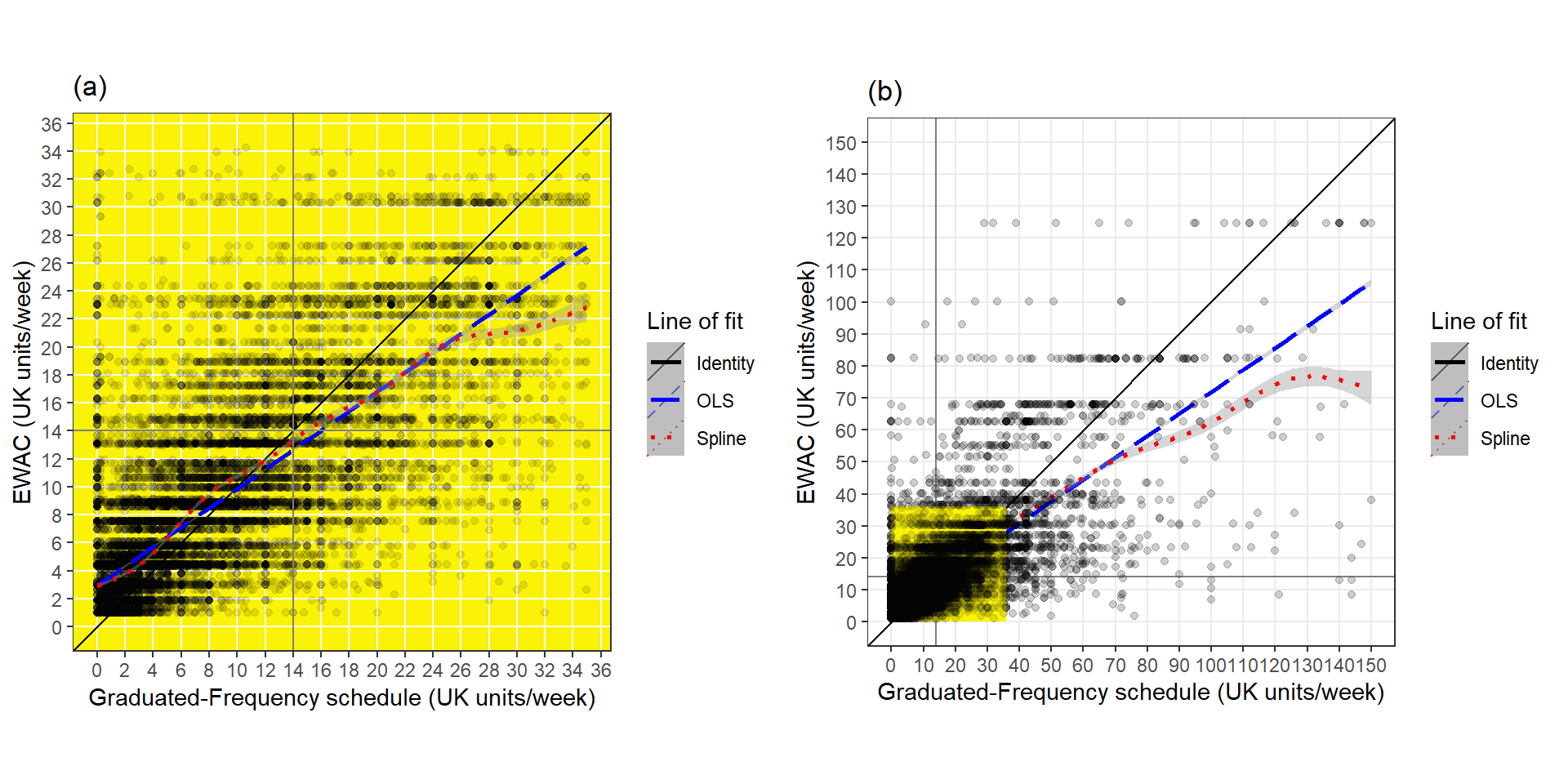


Figure 1: Plots of EWAC against GF in (a) low/increasing risk ATS respondents (n=21,338) and (b) all ATS respondents (n=22,373)

### 3.1.2 Subgroup accuracy

Next, the MD and RMSD are regressed against respondents characteristics in order to identify subgroups with pronounced bias or total error (see supplementary materials 3, Table 7). The model predictors explain a very modest proportion of both MD and RMSD (\(R^2\) statistics < 2%). Nevertheless, specific subgroups do exhibit very different MD and RMSD.

Figure 6 summarises the MD for a selection of subgroups whose predicted MD was either over 1 or under -1; and whose coefficients had a \(p\)-value under 0.05. The reference category’s key characteristics are: females aged between 25 and 34 years of White ethnicity without educational qualifications, whose reported favourite drink was beer. 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. Respondents aged between 55 and 64 years or 75 years or more respectively had MDs 2.2 units [0.5; 3.9] and 4.2 units [0.9; 7.6] in excess of the reference MD. Evidence of milder–this time negative–bias was also found in respondents reporting any religion other than a Christian or Muslim faith.

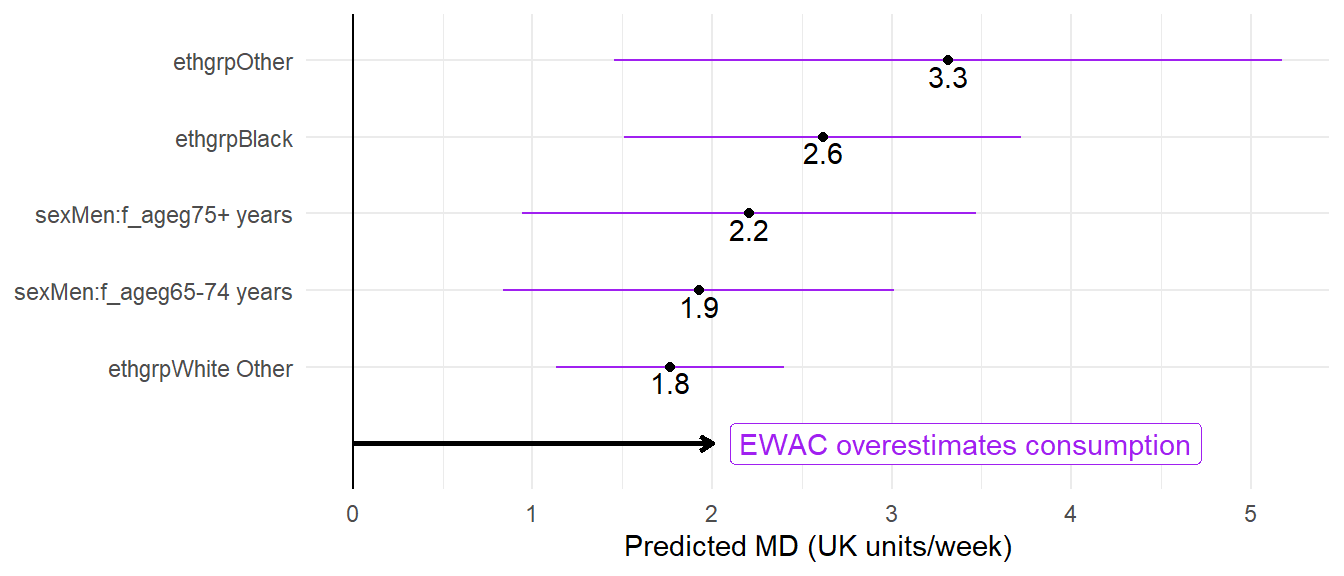


Figure 2: Forest plot of modelled MD for selected subgroups

Similarly, we sought to identify subgroups in which agreement between the EWAC and the GF gold standard was substantially different. Figure 3 shows those subgroups with an RMSD found to be significantly different from the rest of the population (\(p < 0.05\)); and estimated to be 20% greater or smaller than the RMSD of the reference category.

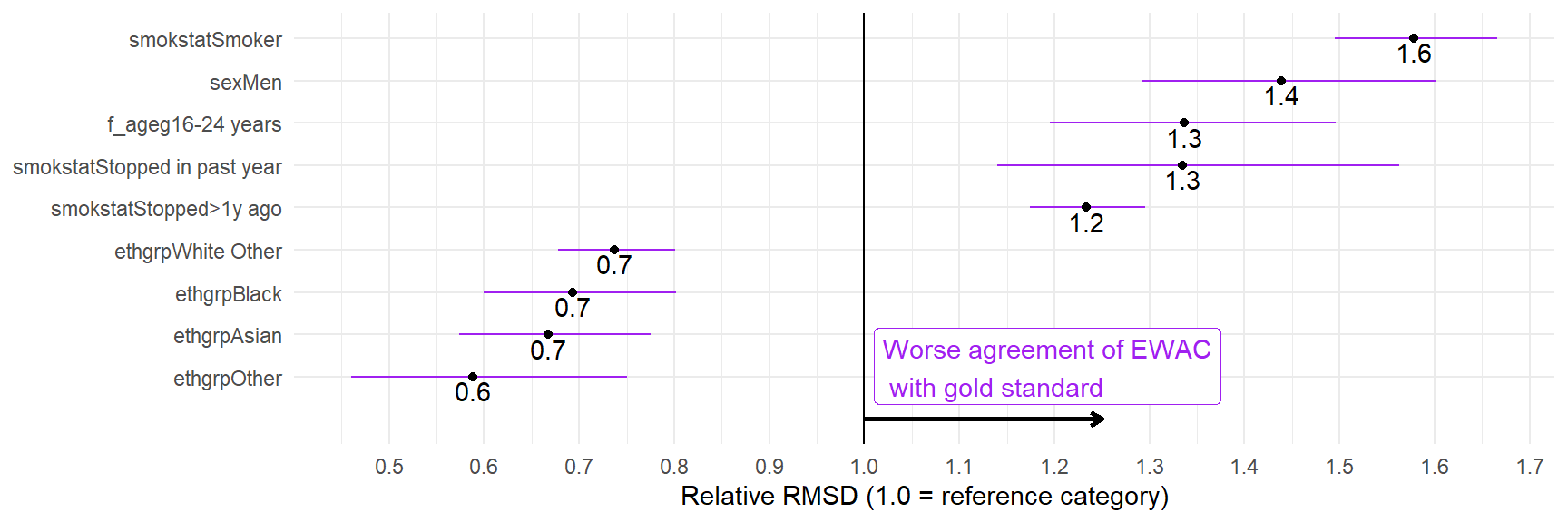
Error is significantly greater in mean (those reporting any religion other than Christian or Muslim; and in current smoker compared to never-smokers.

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

In addition, respondents who attempted to reduce their alcohol consumption in the last year also had a higher RMSD. On the contrary, educational qualifications seem to significantly improve the agreement between EWAC and the gold standard. The RMSD is reduced by an average 2 units with an educational qualification compared to no qualifications (and without effect on MD), suggesting that respondents may have better recall and clarity over alcohol beverage content.

### 3.1.3 Receiver Operating Characteristics

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

The last step of this analysis was to examine the EWAC’s ability to classify participants’ consumption as being equal or in excess of 14 UK units/week. In comparison with the AUDIT-C score, EWAC increases the area under the curve by 5 percentage points (Table 4), a statistically significant improvement (\(p\) < 0.001). ROC curve are available from Supplementary Material 3. The cut-off that maximises the sum of specificity and sensitivity is 9.95 units/week. In constrast, using the nominal cut-off of 14 units/week yields a sensitivity of 0.705 and a specificity of 0.922.

In summary, using an EWAC threshold of 10 units or more over an AUDIT-C score of 5 or more provides equivalent sensitivity, but with an a steep gain in specificity of 13 percentage points.

Table 4: Receiver operating characteristics of AUDIT-C score and EWAC for consumption >= 14 UK units/week (n = 22,404)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Metric | AUC | 95% CI | Best threshold | Sensitivity | Specificity |
| AUDIT-C score | 0.871 | [0.866; 0.876] | 4.5 | 0.882 | 0.684 |
| EWAC | 0.921 | [0.917; 0.925] | 9.95 | 0.876 | 0.816 |

## 3.2 Study 2: Concurrent validity in hospital patients

This study collected TLFB daily alcohol consumption from a total of 105 participants, of whom 63 were classified as low-risk drinkers, 25 as increasing-risk alcohol users (AUDIT-C score of 5 to 7), and 17 as high-risk alcohol users (AUDIT-C score of 8 or more). A total of 13 participants did not provide TLFB information for at least one day missing out of 28, resulting in 5% (137/2940) of TLFB days were missing.

Using only 103 participants with a minimum of 7 days recorded on the TLFB, MD is estimated at -0.5 unit/week, which does not provide any evidence of bias greater than 1 unit/week (\(p\) = 0.316). As for error, RMSD is estimated at 5.4 [95%CI: 3.5; 6.8] and is statistically significantly greater than 2 units (\(p\) < 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.

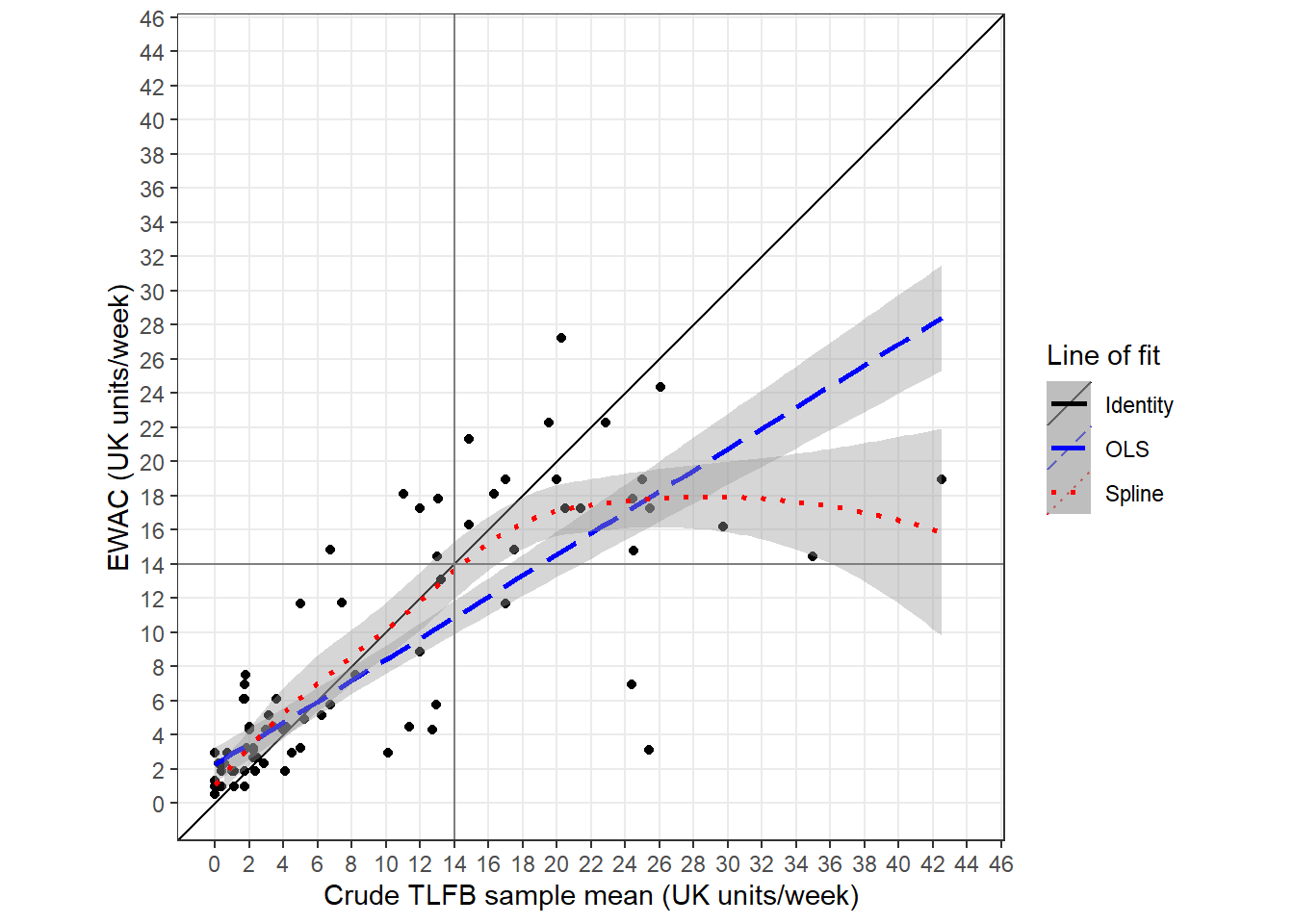


Figure 4: Plot of EWAC estimates against TLFB estimates in hospital participants with lines of fit

## 3.3 Study 3: Empirical distribution functions

Table 6 provides estimates of total alcohol consumption in adult resident of 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 proposed EWAC estimates of total consumption are just 77% of a very reliable estimate, the HSE prospective diary.

Figure 5 suggests that the EWAC, like the ATS GF it was estimated against, estimates a higher prevalence of low risk consumption (<= 14 units/week) and increased risk consumption than HSE. In constrast 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.

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 | 72.4 |
| HSE prospective diary | 12.7 | 8.0 | 264.7 | 4,640 | 66.0 |
| ATS GF | 10.0 | 5.1 | 242.0 | 22,136 | 51.9 |
| ATS EWAC | 9.8 | 4.4 | 245.0 | 25,882 | 50.9 |
| Retail sales | 19.3 | – | – | – | – |

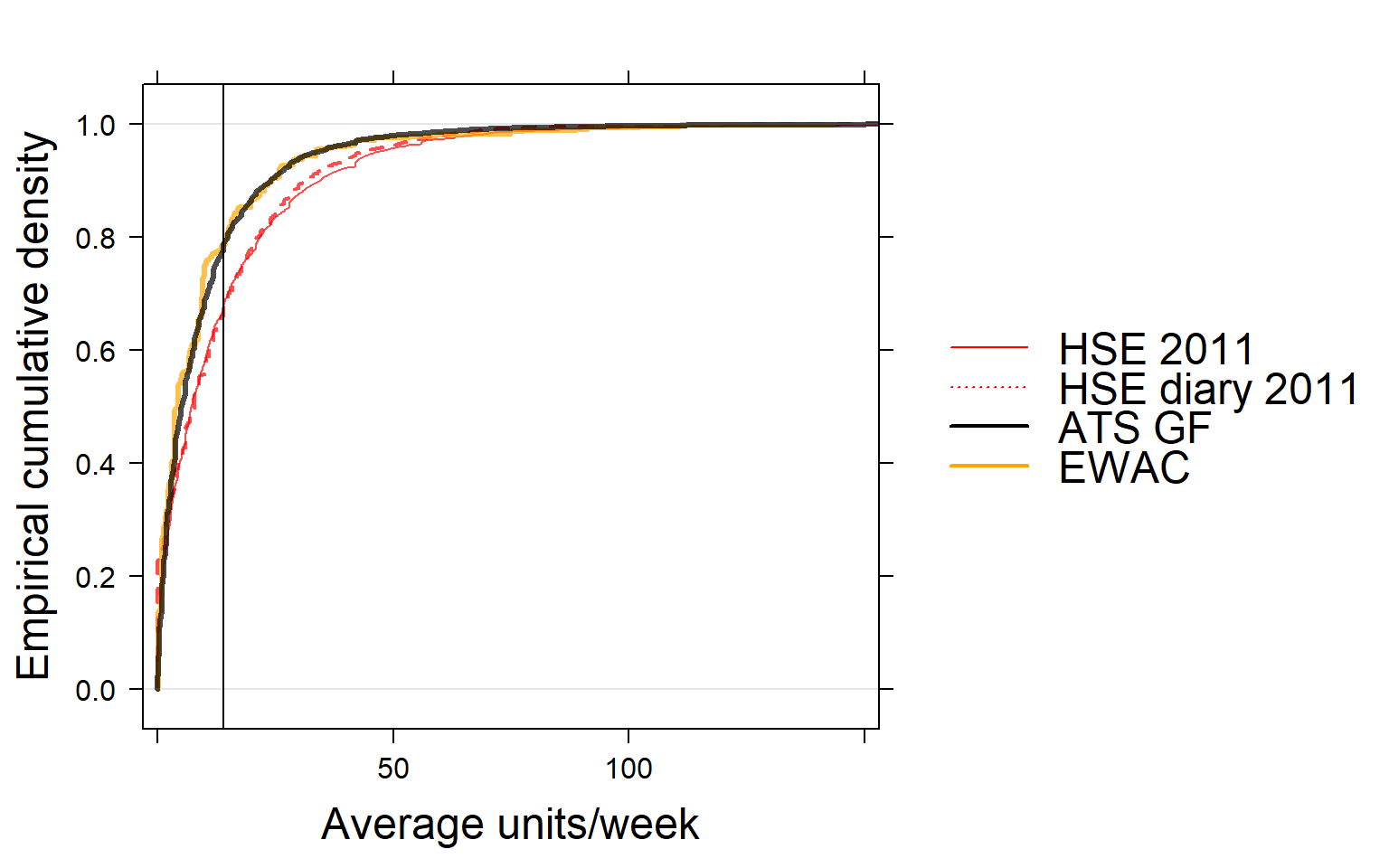


Figure 5: Empirical cumulative distribution function of weekly alcohol consumption in England according to four alcohol schedules in residents aged 18 years and over

# 6 Ethics

Analysis of HSE and ATS microdata was approved by the University of Southampton’s Faculty of Medicine Ethics Committee (ERGO 44682). The hospital questionnaire (Study 3) collection and analysis was approved by the Health Research Authority National Research Ethics Service (IRAS 247458; REC 18/SC/0564).

# 7 Supplementary materials

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

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

## 7.3 Supplementary materials 3: Subgroup analyses

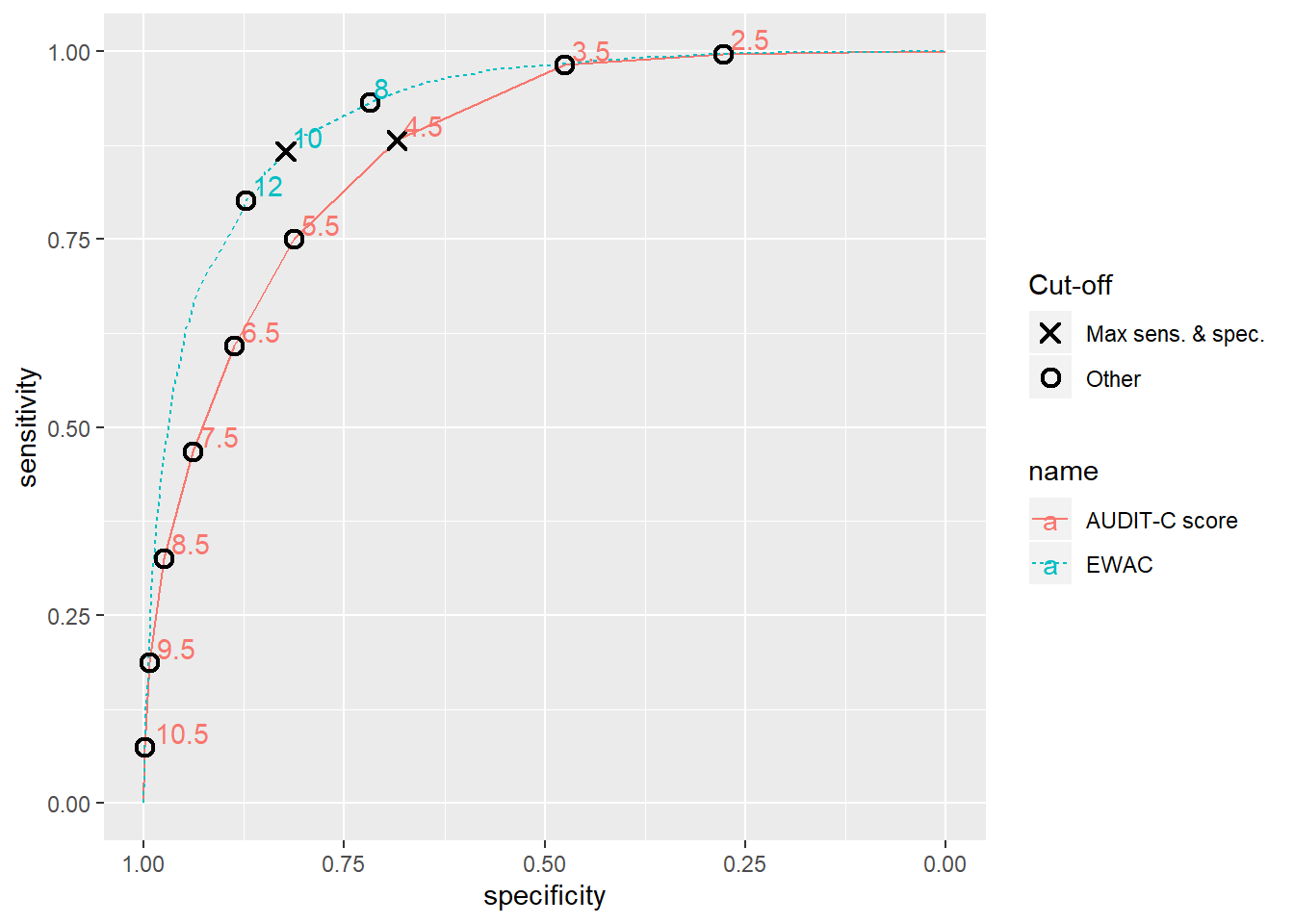
Table 7: Coefficients of linear regression of the bias and error of EWAC compared with GF in all respondents (n = 21,874)

|  |  |  |
| --- | --- | --- |
|  | | |
|  | (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 8: 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)\*\* |
| 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) |
| 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)\*\* |
| 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 | |

## 7.4 Supplementary materials 3: ROC curves



# 9 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] 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.

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

[10] 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).

[11] 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).

[12] 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).

[13] 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).

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

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

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

[17] Lumley T. Survey: Analysis of complex survey samples 2019.

[18] Lumley T. Analysis of complex survey samples. Journal of Statistical Software 2004;9:1–19.

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

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

[21] 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).

[22] 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).

[23] 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).

[24] 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).

[25] 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).

[26] 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).

[27] 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).

[28] 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).

[29] 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.

[30] 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).

[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](https://doi.org/10.17605/OSF.IO/7WE4M).

[32] 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).

[33] Dixon WJ, Massey FJJ. Introduction to statistical analysis. 4th ed. New York: McGraw–Hill; 1983.