Juan Ángel Bellón, Juan de Dios Luna, Michael King, Irwin Nazareth, Emma Motrico, María Josefa GildeGómez-Barragán, Francisco Torres-González, Carmen Montón-Franco, Marta Sánchez-Celaya, Miguel Ángel Díaz-Barreiros, Catalina Vicens and Patricia Moreno-Peral

Predicting the onset of hazardous alcohol drinking in primary care:

development and validation of a simple risk algorithm

Abstract

Background
Little is known about the risk of progressing to hazardous alcohol use in abstinent or low-risk drinkers.

Aim
To develop and validate a simple brief risk algorithm for the onset of hazardous alcohol drinking (HAD) over 12 months for use in primary care.

Design and setting
Prospective cohort study in 32 health centres from six Spanish provinces, with evaluations at baseline, 6 months, and 12 months.

Method
Forty-one risk factors were measured and multilevel logistic regression and inverse probability weighting were used to build the risk algorithm. The outcome was new occurrence of HAD during the study, as measured by the AUDIT.

Results
From the lists of 174 GPs, 3954 adult abstinent or low-risk drinkers were recruited. The 'predictAL-10' risk algorithm included just nine variables (10 questions): province, sex, age, cigarette consumption, perception of financial strain, having ever received treatment for an alcohol problem, childhood sexual abuse, AUDIT-C, and interaction AUDIT-C*Age. The c-index was 0.886 (95% CI = 0.854 to 0.918). The optimal cutoff had a sensitivity of 0.83 and specificity of 0.80. Excluding childhood sexual abuse from the model (the 'predictAL-9'), the c-index was 0.880 (95% CI = 0.847 to 0.913), sensitivity 0.79, and specificity 0.81. There was no statistically significant difference between the c-indexes of predictAL-10 and predictAL-9.

Conclusion
The predictAL-10/9 is a simple and internally valid risk algorithm to predict the onset of hazardous alcohol drinking over 12 months in primary care attendees; it is a brief tool that is potentially useful for primary prevention of hazardous alcohol drinking.

Keywords
alcohol consumption; clinical prediction rule; primary health care.

INTRODUCTION

Alcohol use occupies fifth place among risk factors contributing to worldwide global disease burden.[1] In the European Union 11.8% of all deaths between the ages of 15 and 64 years can be attributed to alcohol,[2] and the absolute risk of dying from an adverse alcohol-related condition increases linearly with the amount of alcohol consumed over a lifetime, with no safe level.[3] Individuals who misuse alcohol (and their families) suffer from physical, mental, and social harm. Apart from being a drug of dependence, for many years alcohol has been known to cause some 60 different types of disease and condition.[2]

Adult per capita alcohol consumption is about 6.2 litres/year on average worldwide, and about 9.6 litres in the high-income countries (the UK 11.6, the US 9.2, Australia 12.2, and Spain 11.2 litres).[4] In Spain the recommended low-risk consumption for adult healthy males is <170 g alcohol per week or <110 g for females.[5] Unhealthy alcohol use includes the full spectrum, from hazardous use to alcohol dependence.[6] Hazardous drinking is defined as 'consumption levels that increase the risk for health consequences' and harmful drinking as 'that which is already causing damage to health (physical or mental)',[7] whereas alcohol abuse and dependence lead to clinically significant impairment or distress (Appendix 1). Around 30% of the population in the US,[6] 24–25% in Canada[8] and the UK,[9] or 18% in Spain[10] are susceptible to risk or harm from their drinking behaviour.

There is widespread knowledge on screening and interventions in hazardous and dependent drinkers,[7,11–13] although questions on the long-term effectiveness of brief interventions for alcohol remain unanswered.[14] Much less is known about the risk of progressing to hazardous use in abstinent or currently low-risk drinkers.[15] Many risk factors are associated with the onset of hazardous or harmful alcohol drinking,[16–19] but so far only one risk algorithm taking into account their

JÁ Bellón, MD, PhD, GP, Centro de Salud El Palo, researcher, Instituto de Investigación Biomédica de Málaga (IBIMA), associate professor, Departamento de Medicina Preventiva y Psiquiatría, Universidad de Málaga, Málaga, Spain. **JD Luna**, PhD, professor, Departamento de Bioestadística; **F Torres-González**, MD, PhD, professor and researcher, Centro de Investigación Biomédica en Red de Salud Mental, Universidad de Granada, Granada, Spain. **M King**, MD, PhD, emeritus professor, Division of Psychiatry, Faculty of Brain Sciences; **I Nazareth**, MD, PhD, professor and head, Department of Primary Care and Population Health, University College London, London, UK. **E Motrico**, PhD, professor, Departamento de Psicología, Universidad Loyola Andalucía, Sevilla, Spain. **MJ GildeGómez-Barragán**, MD, PhD, professor and researcher, Unidad Docente de Medicina Familiar y Comunitaria de La Rioja, La Rioja, Spain. **C Montón-Franco**, MD, PhD, GP, Centro de Salud Casablanca, associate professor, Departamento de Medicina y Psiquiatría, Universidad de Zaragoza, Zaragoza, Spain. **M Sánchez-Celaya**, MD, PhD, director, Continuidad Asistencial del Hospital Universitario Infanta Sofía, Madrid, Spain. **MÁ Díaz-Barreiros**, MD, PhD, GP, Centro de Salud Vecindario, Servicio Canario de Salud, Las Palmas, Spain. **C Vicens**, MD, PhD, GP, Centro de Salud son Serra-La Vileta, Instituto Balear de la Salud, Palma de Mallorca, Illes Balears, Spain. **P Moreno-Peral**, PhD, researcher, Instituto de Investigación Biomédica de Málaga (IBIMA), Málaga, Spain.

Address for correspondence
Juan Ángel Bellón, Departamento de Medicina Preventiva, Salud Pública y Psiquiatría, Facultad de Medicina, Universidad de Málaga, Campus de Teatinos, 29071 Málaga, Spain.

E-mail: jabellon@uma.es

Submitted: 20 June 2016; **Editor's response:** 26 August 2016; **final acceptance:** 4 November 2016.

How this fits in

Little is known about the risk of progressing to hazardous alcohol use in people who are abstinent or low-risk drinkers. Only one risk algorithm is currently available in primary care to predict the onset of hazardous alcohol drinking (the predictAL), with the patient having to answer 29 questions to calculate their risk at 6 months. The predictAL-10/9 risk algorithm is a shorter alternative (9–10 questions), which has higher discriminative validity and allows longer-term predictions (12 months).

combined effect has been published; the predictAL.[15] This algorithm, which was internally validated in six European countries and externally validated in Chile, predicts the onset of hazardous alcohol drinking (HAD) at 6 months in primary care attendees. The predictAL has good discriminative validity, but to obtain a risk probability for a particular patient requires administering two questionnaires, the AUDIT (10 items) and the PRIME-MD-Anxiety (Panic-Syndrome) (15 items), as well as another four items (sex, age, country, and lifetime alcohol problem). Accordingly, time management might be a barrier to its use, given the competing demands in busy clinical practice settings.[20] Moreover, a prediction period beyond 6 months may also be useful because a relatively high proportion of abstinent or low-risk drinkers will develop HAD at 12 months but not at 6 months.[21] Therefore, this study aimed at developing and internally validating a shorter and simpler risk algorithm to predict the onset of HAD over 12 months in primary care.

METHOD

Design and setting

A prospective cohort study was undertaken with evaluations at baseline, 6 months, and 12 months. Although this cohort was originally recruited with the aim of developing a risk model for the onset of major depression,[22] this analysis aimed at predicting the onset of HAD. The method has been described in detail elsewhere[22] and the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) was followed.[23] The predictAL-Spain study was conducted with the participation of 174 GPs belonging to 32 health centres (mean 5.4, range 1–10) distributed throughout Spain (in six provinces). Each health centre covers a population of 15 000–30 000 inhabitants from a geographically defined area. The GPs in each health centre work as a group, with extensive primary care teams. The Spanish National Health Service provides free medical cover at the point of access to 95% of the population. Patients can visit their GP as often as they wish without having to pay for it, even when they do so for preventive reasons. Each patient is assigned to only one GP, who has gatekeeper functions. The health centres taking part cover urban and rural settings in each province.

Sampling and exclusion criteria

Random samples of four to six attendees from GP appointment lists were taken for each day of recruitment. The GPs introduced the study to the selected patients, checked their exclusion criteria, and requested their permission before contacting the researcher. Participants who gave informed consent undertook a research interview given by research assistants within 2 weeks. The study population was recruited between October 2005 and February 2006. Exclusion criteria were an inability to understand or speak Spanish, severe mental disorder (for example, psychosis, bipolar disorder), dementia or severe neurological/sensory illness, terminal illness, the person was scheduled to be out of the city for more than 3 months during the 12 months of follow-up, and persons (representatives) who attended the GP's office on behalf of the person who had the appointment.

Variables

Outcome measure. The outcome was new occurrence of HAD during the 12-month study. Alcohol use in the preceding 6 months was assessed at 6 and 12 months of follow-up by the Alcohol Use Disorders Identification Test (AUDIT).[24] The AUDIT is a 10-item questionnaire that addresses frequency of alcohol consumption, alcohol-related problems, and alcohol dependence symptoms. It was specifically developed for use in a primary care population and has good validity and reliability in many countries, including Spain.[24]

To classify a person as a hazardous alcohol drinker, an AUDIT cutoff of ≥8 for males and ≥6 for females was used. Other Spanish researchers indicate that a cutoff of ≥8 has a sensitivity and specificity for females and males together of 0.90,[25] and a cutoff of ≥6 for females has a sensitivity of 0.90 and a specificity of 0.95.[26] Cronbach's α varies between 0.86[25] and 0.93,[26] and test–retest reliability (intraclass correlation coefficient) is 0.90.[25]

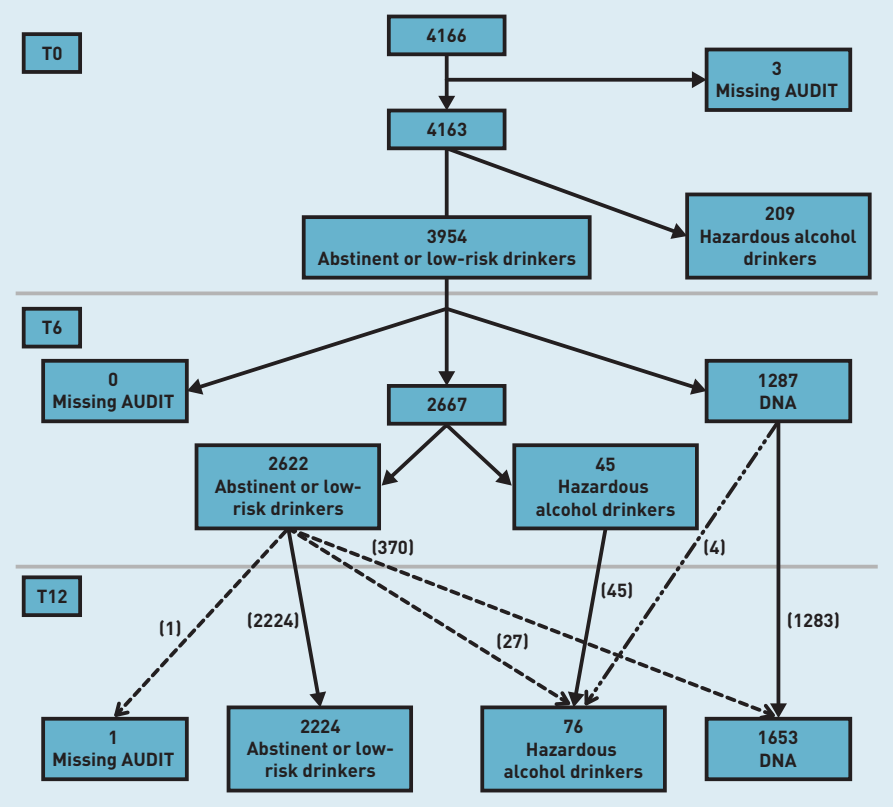


Figure 1. Flowchart of patients through the predictAL-10/9 study and numbers becoming hazardous alcohol drinkers. AUDIT (Alcohol Use Disorders Identification Test). Hazardous alcohol drinking (males: AUDIT≥8; females: AUDIT≥6). Abstinent or low-risk drinkers (males: AUDIT<8; females: AUDIT<6); DNA = did not attend; T0 = at baseline; T6 = at 6 months; T12 = at 12 months.

Measurement of potential risk factors. Forty-one potential risk factors were selected for which there was evidence of reliability and validity in the questionnaires used to evaluate them.[22] Baseline measurements were made of all the potential risk factors by independent research assistants who were blind to the objective of the study. All risk factors are described in detail elsewhere,[22] and a summary is given in Appendix 1.

Statistical analysis

Participants who did not complete the AUDIT at both 6 and 12 months were excluded, and those who were hazardous alcohol drinkers at either or both (6 or 12 months) were considered to be hazardous alcohol drinkers in the outcome. Multilevel logistic regression was performed including health centre as a random component (Appendix 1). The required sample size was estimated based on the need for at least 10 outcome events (HAD) per independent variable included in the prediction rule.[27]

Variables were selected using a threshold for inclusion of $P<0.20$ to ensure that information lost as a result of exclusion of a variable from the equation was minimal.[28] From the model thus obtained, those variables with $P>0.05$ were extracted step by step to obtain a more parsimonious model. Pair-wise interactions between the variables in the model and sex and age were tested. Inverse probability weighting[29,30] was used to adjust for a possible attrition bias because of participants lost to follow-up.

The c-index[31] was calculated to estimate the discriminative validity of the final predictAL-10/9 models. To compare the discriminative validity between risk algorithms, the test was performed for correlated c-indexes. Prediction models derived with multivariable regression analysis are known to overestimate regression coefficients. A calculation proposed by Copas[32] was used to estimate overfitting of the prediction models. Effect sizes were calculated using Hedges' g[33] for the difference in log odds of the predicted probability between patients who were later observed to be hazardous drinkers and those who were not. Calibration, which is the agreement between the observed proportions of HAD and the predicted risks, was studied with calibration plots taking deciles of risk.

Finally, the optimal threshold values (cutoff points) where Youden's J statistic (J = Sensitivity + Specificity – 1)[34] was greater were highlighted. All analyses were conducted using Stata (version 13.1). All reported *P*-values were two-sided.

RESULTS

Of the 6299 primary care attendees approached, 1251 (19.9%) were excluded: 506 (8.03%) were outside the age range (18–75 years); 446 (7.1%) were either representatives of patients or did not attend the appointment; 156 (2.5%) had a severe mental disorder, dementia, or severe neurological/sensory illness; 63 (1.0%) terminal illness; 47 (0.75%) trouble communicating in Spanish; and 33 (0.52%) were scheduled to be out of the city for longer than 3 months during the 12 months of follow-up. Of the remaining 5048 patients asked to take part in the study 4166 (82.5%) gave their consent. These were then interviewed at baseline, but 209 (5.02%) were hazardous alcohol drinkers (by AUDIT) and three (0.07%) had a missing diagnosis, so they were also excluded. Thus, the at-risk population comprised 3954 patients (Figure 1). The patients' sociodemographic characteristics are shown in Table 1. Of the 3954 patients, 2667 (67.5%) were interviewed at 6 months and 2301 (58.2%) at 12 months. The main baseline variables associated with drop-outs were province (Majorca and Las Palmas), sex (male), lower age, country of birth (outside Spain), lower educational level, never having enough money to afford

Table 1. Demographic characteristics of abstinent or low-risk drinkers (population at risk)

Demographic characteristics	Granada	Zaragoza	Madrid	La Rioja	Majorca	Las Palmas	Total
Abstinent or low-risk drinkers, *n* (%)	731 (18.5)	715 (18.1)	724 (18.3)	727 (18.4)	695 (17.6)	362 (9.2)	3954 (100)
Age, years, mean (SD)	49.66 (16.12)	46.88 (15.44)	50.41 (15.67)	49.29 (15.5)	49.61 (15.65)	43.94 (14.29)	48.7 (15.66)
Sex, *n* (%)							
Female	552 (75.51)	475 (66.43)	507 (70.03)	482 (66.3)	465 (66.91)	272 (75.14)	2753 (69.63)
Male	179 (24.49)	240 (33.57)	217 (29.97)	245 (33.7)	230 (33.09)	90 (24.86)	1201 (30.37)
Marital status, *n* (%)							
Married	495 (67.72)	467 (65.31)	490 (67.68)	477 (65.61)	438 (63.02)	183 (50.55)	2550 (64.49)
Separated	35 (4.79)	23 (3.22)	36 (4.97)	30 (4.13)	41 (5.9)	35 (9.67)	200 (5.06)
Divorced	6 (0.82)	13 (1.82)	20 (2.76)	9 (1.24)	24 (3.45)	19 (5.25)	91 (2.3)
Single	128 (17.51)	176 (24.62)	136 (18.78)	163 (22.42)	137 (19.71)	101 (27.9)	841 (21.27)
Widowed	67 (9.17)	36 (5.03)	42 (5.80)	48 (6.6)	55 (7.91)	22 (6.08)	270 (6.83)
Missing	0	0	0	0	0	2 (0.55)	2 (0.05)
Household status, *n* (%)							
Not living alone	663 (90.7)	666 (93.15)	665 (91.85)	660 (90.78)	608 (87.48)	342 (94.48)	3604 (91.15)
Living alone	68 (9.3)	49 (6.85)	59 (8.15)	67 (9.22)	87 (12.52)	20 (5.52)	350 (8.85)
Education, *n* (%)							
Higher education	84 (11.49)	109 (15.24)	67 (9.25)	119 (16.37)	40 (5.76)	44 (12.15)	463 (11.71)
Secondary	127 (17.37)	182 (25.45)	171 (23.62)	151 (20.77)	118 (16.98)	90 (24.86)	839 (21.22)
Primary	273 (37.35)	344 (48.11)	287 (39.64)	404 (55.57)	417 (60.0)	159 (43.92)	1884 (47.65)
Trade/other	247 (33.79)	80 (11.19)	199 (27.49)	52 (7.15)	120 (17.27)	69 (19.06)	767 (19.4)
Missing	0	0	0	1 (0.14)	0	0	1 (0.03)
Employment, *n* (%)							
Employed	244 (33.38)	364 (50.91)	325 (44.89)	360 (49.52)	262 (37.7)	197 (54.42)	1752 (44.31)
Unemployed	52 (7.11)	44 (6.15)	35 (4.83)	49 (6.74)	44 (6.33)	44 (12.15)	268 (6.78)
Retired	147 (20.11)	112 (15.66)	148 (20.44)	155 (21.32)	134 (19.28)	30 (8.29)	726 (18.36)
Unable to work	68 (9.3)	16 (2.24)	41 (5.66)	7 (0.96)	133 (19.14)	17 (4.7)	282 (7.13)
Looking after family	191 (26.13)	150 (20.98)	163 (22.51)	137 (18.84)	116 (16.69)	63 (17.4)	820 (20.74)
Full-time student	26 (3.56)	27 (3.78)	9 (1.24)	18 (2.48)	5 (0.72)	7 (1.93)	92 (2.33)
Other	1 (0.14)	1 (0.14)	1 (0.14)	1 (0.14)	1 (0.14)	4 (1.1)	9 (0.23)
Missing	2 (0.27)	1 (0.14)	2 (0.28)	0	0	0	5 (0.13)
Country of birth, *n* (%)							
Spain	711 (97.26)	683 (95.52)	672 (92.82)	687 (94.5)	645 (92.81)	321 (88.67)	3719 (94.06)
Other	19 (2.6)	32 (4.48)	44 (6.08)	40 (5.5)	43 (6.19)	34 (9.39)	212 (5.36)
Missing	1 (0.14)	0	8 (1.1)	0	7 (1.01)	7 (1.93)	23 (0.58)
Ethnic group, *n* (%)							
White European	713 (97.54)	599 (83.78)	698 (96.41)	687 (94.5)	678 (97.55)	360 (99.45)	3735 (94.46)
Other ethnic group	12 (1.64)	6 (0.84)	24 (3.31)	24 (3.3)	12 (1.73)	2 (0.55)	80 (2.02)
Missing	6 (0.82)	110 (15.38)	2 (0.28)	16 (2.2)	5 (0.72)	0	139 (3.52)
Financial strain, *n* (%)							
Living comfortably	55 (7.52)	81 (11.33)	42 (5.82)	61 (8.4)	56 (8.06)	19 (5.26)	314 (7.95)
Doing all right	496 (67.85)	532 (74.41)	506 (70.08)	586 (80.72)	480 (69.06)	252 (69.81)	2852 (72.2)
Finding it difficult or very difficult	180 (24.62)	102 (14.27)	174 (24.1)	79 (10.88)	159 (22.88)	90 (24.93)	784 (19.85)
Missing	0	0	2 (0.28)	1 (0.14)	0	1 (0.28)	4 (0.1)

food or clothing (basic financial strain), housing status (rented), and dissatisfaction with the area where they lived (Appendix 2). The AUDIT score at baseline was not a predictor of drop-outs.

Forty-five of those successfully contacted at 6 months had become engaged in HAD and a further 31 of those contacted at 12 months had become engaged in HAD. All those who were engaged in HAD at 6 months were still engaged in HAD at 12 months (Figure 1). The final model, the predictAL-10, to predict the onset of HAD at 12 months in primary care attendees included nine variables (10 questions) (Table 2): province, sex (male), age (lower), cigarette consumption, perceived financial strain, having ever received treatment for an alcohol problem, childhood sexual abuse, AUDIT-C, and the interaction AUDIT-C*Age. The shrinkage factor was 0.9595 (shrinkage = 1 indicates that there is no overestimation). The c-index was 0.886 (95% CI = 0.854 to 0.918) and the effect size (Hedges' g) 1.694 (95% CI = 1.460 to 1.928). The calibration showed an accurate

Table 2. The predictAL-10 and predictAL-9 models[a] to predict the onset of hazardous alcohol drinking at 12 months

	Incidence of hazardous alcohol drinking		PredictAL-10[b] (*N* = 2264)				PredictAL-9[c] (*N* = 2278) (excluding childhood sexual abuse from the model)			
Risk factors	**No**	**Yes (%)**	**OR**	**OR[d]**	**95% CI**	***P***	**OR**	**OR[e]**	**95% CI**	***P*-value**
Constant			0.0008	0.0009	0.0001 to 0.0071	<0.001	0.0011	0.0012	0.0001 to 0.0079	<0.001
Province										
Granada (Reference)	499	7 (1.38)	1.0	1.0		1.0	1.0			
Zaragoza	447	16 (3.46)	2.02	1.94	0.49 to 8.37	0.333	1.80	1.73	0.49 to 6.70	0.379
Madrid	406	4 (0.98)	0.72	0.69	0.15 to 3.56	0.690	0.66	0.64	0.15 to 2.93	0.585
Logroño (La Rioja)	429	29 (6.33)	7.12	6.84	2.05 to 24.79	0.002	6.46	6.22	2.11 to 19.79	0.001
Majorca	258	13 (4.80)	5.32	5.11	1.11 to 25.62	0.037	4.74	4.57	1.08 to 20.84	0.040
Las Palmas	185	7 (3.65)	3.16	3.03	0.61 to 16.28	0.170	3.09	2.96	0.62 to 15.41	0.168
Sex										
Female (Reference)	1622	28 (1.70)	1.0	1.0		1.0	1.0			
Male	602	48 (7.38)	3.20	3.07	1.29 to 7.91	0.012	2.82	2.72	1.62 to 6.83	0.022
Age (range 18–75 years)			0.993	0.953	0.972 to 1.015	0.539	0.989	0.952	0.969 to 1.010	0.316
AUDIT-C			2.51	2.41	1.63 to 3.85	<0.001	2.37	2.28	1.57 to 3.59	<0.001
AUDIT-C*Age[f]			0.991	0.951	0.984 to 0.999	0.045	0.993	0.956	0.985 to 1.000	0.068
Cigarette consumption per day										
Non-smoking (Reference)	1756	41 (2.28)	1.0	1.0		1.0	1.0			
<10	181	12 (6.22)	2.39	2.29	1.21 to 4.73	0.012	2.42	2.33	1.20 to 4.88	0.014
10–20	200	11 (5.21)	1.28	1.23	0.51 to 3.18	0.600	1.38	1.33	0.57 to 3.31	0.475
>20	87	12 (12.1)	3.48	3.34	1.31 to 9.27	0.013	3.59	3.46	1.33 to 9.68	0.012
Financial strain										
Living comfortably (Reference)	183	3 (1.61)	1.0	1.0		1.0	1.0			
Doing all right	1634	56 (3.31)	1.94	1.86	0.48 to 7.82	0.351	2.07	1.99	0.50 to 8.51	0.313
Finding it difficult or very difficult	405	17 (4.03)	4.19	4.02	0.98 to 17.84	0.053	4.66	4.49	1.08 to 20.03	0.039
Ever treated for alcohol problems										
No (Reference)	2204	73 (3.21)	1.0	1.0		1.0	1.0			
Yes	12	3 (20.0)	11.77	11.30	1.98 to 70.05	0.007	13.19	12.70	2.61 to 66.63	0.002
Childhood sexual abuse										
No (never) (Reference)	2161	70 (3.14)	1.0	1.0						
Yes (rarely, sometimes, often, frequently)	49	6 (10.9)	5.07	4.87	1.71 to 15.09	0.003				

[a]Multilevel logistic regression with health centre as a random component and weighting for the inverse probability of remaining in the follow-up to 12 months. [b]Discriminative validity: c-index: 0.886 (95% CI = 0.854 to 0.918) and effect size (Hedges' g): 1.694 (95% CI = 1.460 to 1.928). [c]Discriminative validity: c-index: 0.880 (95% CI = 0.847 to 0.913) and effect size (Hedges' g): 1.658 (95% CI = 1.425 to 1.892). [d]Overfitting estimate: Copas' shrinkage factor = 0.960. [e]Overfitting estimate: Copas' shrinkage factor = 0.963. [f]Likelihood ratio test for the interaction: χ^2 (degree of freedom:1) = 5.84; P *= 0.0157. AUDIT = Alcohol Use Disorders Identification Test.*

goodness-of-fit (Figure 2). The predicted probability (optimal threshold) of 2.72% included 495 (21.9%) primary care attendees who were abstinent or low-risk drinkers at baseline and its sensitivity and specificity were 0.83 and 0.80, respectively (Table 3).

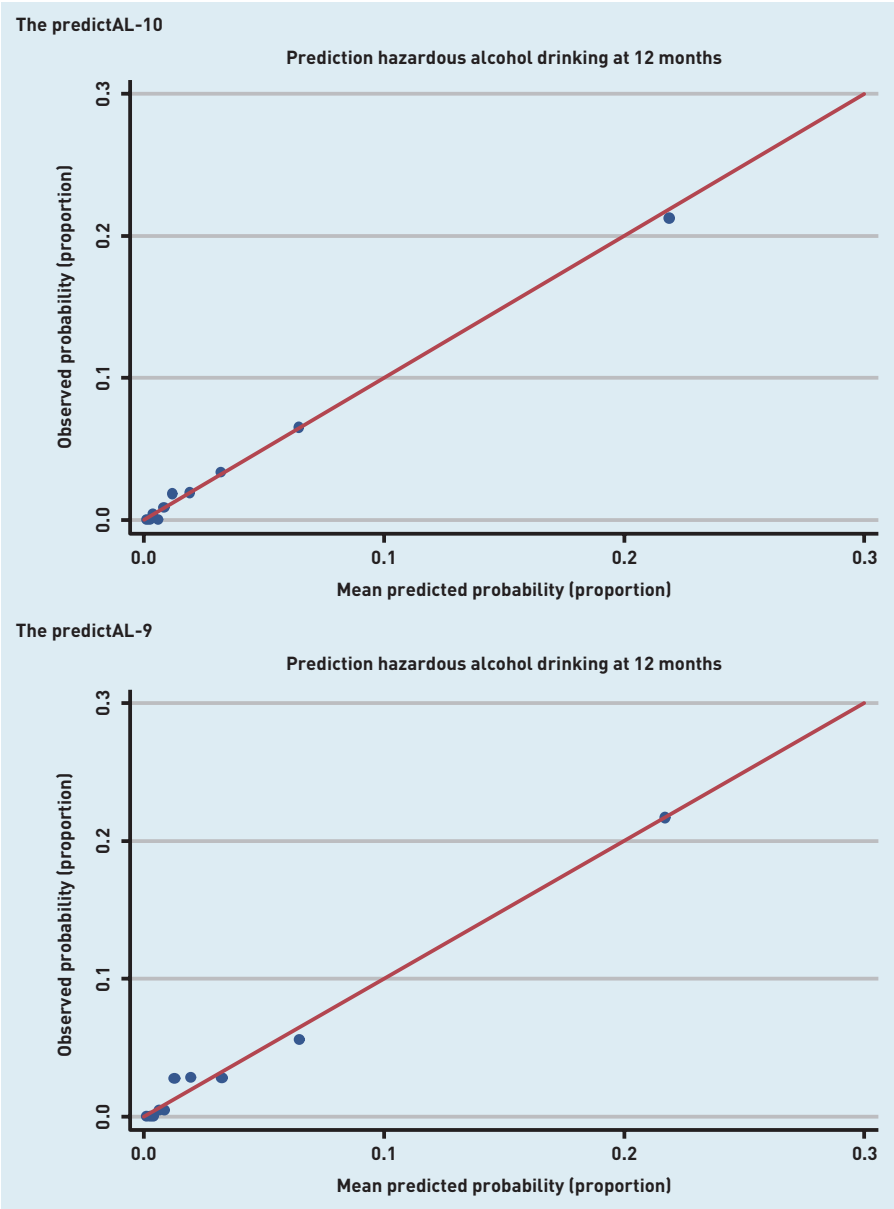
When the potentially sensitive question about childhood sexual abuse was removed from the analysis, the predictAL-9 model was obtained. Compared with the predictAL-10, several coefficients were slightly different (Table 2), sensitivity decreased by 4 points, specificity increased by 1 point, and the calibration plot showed that two deciles of risk were minimally overestimated (Figure 2). The c-indexes, Hedges' g, and shrinkage factors were similar, however, to the predictAL-10 (Table 3). There was no statistically significant difference between the c-indexes of the predictAL-10 and the predictAL-9 (χ^2 1.39; $P = 0.238$). There were 22 (0.96%) and 36 (1.6%) missing values for predictor variables in the predictAL-9 and predictAL-10 models, respectively.

Finally, all the AUDIT questions were entered into the analysis (predictAL-17), but it was found that this did not improve the c-index over the predictAL-10 (χ^2 0.01; $P = 0.941$).

Very slight differences were found between the predictAL-10 and the same model weighted for the inverse probability of remaining in the follow-up to 12 months (Appendix 3). A calculator of the predictAL-9/10 is available on the 'predictplusprevent' website (http://www.predictplusprevent.com/indexDefi.php?idioma=en).

This includes a set of calculators to predict the occurrence of future episodes

Figure 2. Calibration plots (mean predicted probability against observed probability of hazardous alcohol drinking within deciles of predicted risk) of the predictAL-10/9 risk algorithms.



of depression, anxiety, and/or hazardous alcohol drinking in those persons who are not suffering them currently. The website also provides information on activities and interventions to prevent them.

DISCUSSION

Summary

The predictAL-10/9 is a simple, brief, and internally valid risk algorithm to predict the onset of HAD over 12 months in primary care attendees who are abstinent or low-risk drinkers. The predictive model showed an accurate goodness-of-fit and the level of overfitting was minimal.

Strengths and limitations

Of those abstinent and low-risk drinkers who developed HAD, 59% did so by 6 months and the remainder thereafter. This suggests the need to characterise the population at risk of HAD at 12 months versus 6 months, which is an advantage of the predictAL-9/10 over the predictAL[15] (only available for 6 months). To the authors' knowledge, the predictAL-10/9 is the first risk algorithm to predict the onset of HAD over 12 months in primary care. The predictAL-9/10 also had higher discriminative validity than the predictAL[15] and other risk algorithms for the onset of major depression[35–38] and anxiety syndromes[39,40] in primary care as well as for risk indices for cardiovascular events.[41,42]

The study has some limitations. The sample size was not large enough to

Table 3. Discriminative validity of the clinical rule predictions for the onset of hazardous alcohol drinking over 12 months in primary care

Risk algorithms	Number of items	*N*	C-index (95% CI)	Hedges' g (95% CI)	Predicted probability[a]	Frequency[b] *N* (%)	Sensitivity Specificity	LR+ LR–	PPV NPV	Shrinkage factor[c]
predictAL-10[d]	10	2264	0.886 (0.854 to 0.918)	1.694 (1.460 to 1.928)	≥2.72%	495 (21.86)	0.83 0.80	4.15 0.21	0.13 0.99	0.9595
predictAL-17[e]	17	2264	0.886 (0.853 to 0.919)	1.729 (1.495 to 1.963)	≥2.92%	454 (20.05)	0.80 0.82	4.44 0.25	0.13 0.99	0.9595
predictAL-7[f]	7	2264	0.819 (0.772 to 0.866)	1.292 (1.052 to 1.532)	≥3.31%	539 (23.81)	0.71 0.79	3.38 0.37	0.10 0.99	0.8615
Clinical rule predictions excluding the variable sexual abuses in childhood										
predictAL-9[d]	9	2278	0.880 (0.847 to 0.913)	1.658 (1.425 to 1.892)	≥3.01%	475 (20.85)	0.79 0.81	4.16 0.26	0.13 0.99	0.9629
predictAL-16[e]	16	2278	0.883 (0.850 to 0.916)	1.697 (1.463 to 1.931)	≥2.78%	483 (21.20)	0.80 0.81	4.21 0.25	0.13 0.99	0.9639
predictAL-6[f]	6	2278	0.803 (0.752 to 0.854)	1.273 (1.042 to 1.505)	≥2.60%	668 (29.32)	0.70 0.72	2.5 0.42	0.08 0.98	0.8563
Clinical rule predictions including only the AUDIT										
AUDIT-C	3	2288	0.775 (0.721 to 0.830)	1.211 (0.980 to 1.442)	≥2.34%	517 (22.60)	0.75 0.68	4.16 0.26	0.07 0.89	0.9819
AUDIT	10	2288	0.781 (0.725 to 0.836)	1.254 (1.042 to 1.485)	≥2.20%	525 (22.95)	0.75 0.68	4.21 0.25	0.07 0.89	0.9822

[a]Predicted probability of hazardous alcohol drinking at 12 months, cutoff point where Youden's J statistic (J = Sensitivity + Specificity – 1) was greater: 'optimal threshold'. [b]Number of primary care attendees above the optimal threshold. [c]Copas' shrinkage factor estimates overfitting of the prediction models (shrinkage = 1 indicates that there is no overestimation). [d]Including the AUDIT-C (three items). [e]Including the AUDIT (10 items). [f]Excluding any AUDIT. AUDIT = Alcohol Use Disorders Identification Test. LR+ = positive likelihood ratio. LR– = negative likelihood ratio. PPV = positive predictive value. NPV = negative predictive value.

address external validation in this study, such as deriving the algorithm in some provinces and validating it in others. New validations of the predictAL-9/10 are needed in other countries. The study did not include patients <18 years of age; therefore new risk algorithms including young adolescents should be developed and validated because the age of onset for alcohol drinking is between 11 and 15 years, and drinking clearly increases throughout adolescence.[43,44] Relatively few events occurred (76 people developed HAD). When the number of events is low relative to the number of predictors, standard regression could produce overfitted risk models that make inaccurate predictions in other settings.[27] According to the calibration plot (Figure 2) and the Copas shrinkage factor (0.96), however, the level of overfitting was minimal. This study included a large number of GPs and health centres from six provinces in southern, central, and northern Spain, and only a few patients refused to participate. Therefore the sample may be representative of primary care attendees in Spain, although patients who attend infrequently may have been under-represented.[45] Additionally, as hazardous alcohol drinkers visit their GPs less often than low-risk drinkers,[10] the incidence of the onset of HAD may have been underestimated, although estimating the incidence was not the aim of the study. Although there were 41.8% dropouts in the 12 months, only slight differences were seen between the predictAL-10 models with and without inverse probability weighting (Appendix 3), indicating that loss to follow-up was unlikely to lead to attrition bias.[29]

Comparison with existing literature

The nine variables (10 items) included in the predictAL-10 are well-known risk factors for hazardous and harmful alcohol drinking. The variable province had a relevant contribution in predicting the onset of HAD, major depression,[36] and anxiety syndromes.[39] This was also the case for country in international risk algorithms to predict such disorders,[15,35,40] and suggests that geographical variability must be taken into account in the prediction models. It is suggested that an average of the coefficients of the five Spanish provinces is used when the predictAL-10/9 is applied to obtain the probability of the risk of HAD outside Spain. This is the way it is calculated on the 'predictplusprevent' website.

Funding

This study was supported by the Spanish Ministry of Health, the Institute of Health Carlos III, and the European Regional Development Fund 'Una manera de hacer Europa' (grant FIS references: PI041980, PI041771, PI042450, and PI06/1442) and the Andalusian Council of Health (grant references: 05/403 and 06/278); as well as the Spanish Network of Primary Care Research 'redIAPP' (RD06/0018), the 'Aragón group' (RD06/0018/0020), the 'Baleares group' (RD07/0018/0033), and the 'SAMSERAP group' (RD06/0018/0039). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Ethical approval

The predictAL–Spain study was conducted according to the principles expressed in the Declaration of Helsinki. This study complies with the Code of Ethics of the World Medical Association and was approved by the ethics committees: Ethics Committee on Human Research of the University of Granada, Ethics and Research Committee of the Primary Health District of Málaga, Ethics Committee for Clinical Research of Aragon (CEICA). Research assistants explained to patients the predictAL-10 study in detail, their commitments and rights, and answered any questions the patients wished to ask. All participants read an information sheet and signed consent forms to take part in the study.

Provenance

Freely submitted; externally peer reviewed.

Competing interests

The authors have declared no competing interests.

Acknowledgements

The authors thank the Primary Care District of Málaga, particularly Jose Miguel Morales, Javier Navarro, and Maximiliano Vilaseca for their support. Thanks also to the Institute of Health Carlos III (ISCIII), the European Regional Development Fund, the Andalusian Council of Health, the Biomedical Research Institute of Málaga (IBIMA), and the Spanish Network of Primary Care Research (redIAPP) for their economic and logistical support, and we acknowledge the role of the Maristán Network, which facilitated collaboration among a number of the authors on this paper. The authors want especially to thank all the interviewers, patients, and primary care physicians for their participation in the predictAL-10 study.

Discuss this article

Contribute and read comments about this article: **bjgp.org/letters**

The results of this study on sex and age, and HAD, are consistent with the literature. In most countries males tend to drink more than females,[2] and alcohol use occupies third place for males and twelfth for females among risk factors contributing to global disease burden.[1] The incidence of new cases of HAD and alcohol dependence is greater in males aged between 20 and 29 years,[46] and the mean age of the transition from low-risk drinkers to regular HAD is 20 years.[21]

A prediction model containing only the AUDIT or AUDIT-C had a c-index of 0.78 (sensitivity 0.75 and specificity 0.68) to predict the onset of HAD (Table 3), so the predictAL-10/9 (c-index 0.89; sensitivity 0.83 and specificity 0.80) is clearly better than the AUDIT in discriminative validity (Table 3). Lower age and higher AUDIT-C score in these abstinent or low-risk drinkers were associated with an increased incidence of HAD over 12 months. Besides this main effect, an interaction was also found with AUDIT-C*Age, such that older people with higher AUDIT-C scores had a lower risk of developing HAD. In southern Europe the 'Mediterranean' way of drinking, which involves regular, moderate wine consumption mainly with food, increases with age,[47] whereas younger people in Europe generally prefer beer, strong spirits, and binge drinking.[47–49]

Evidence from longitudinal studies suggests that the perception of difficulty managing changes in living arrangements and individual deprivation are associated with HAD.[50] Daily and non-daily smokers are at a greater risk for hazardous drinking and alcohol use disorders.[16,51] Smoking increases the risk for alcohol misuse and it is likely it has a causal role in this relationship.[52]

A biological mechanism underlying the association between alcohol use and smoking has been proposed.[53] Even after being an abstinent or low-risk drinker for at least 6 months, having ever received treatment for an alcohol problem was a strong predictor of future hazardous drinking. There is little doubt about the predictive power of this risk factor.[54]

The three types of child abuse have been associated with alcohol dependence.[17,55] Although childhood sexual abuse could have implications for the course of prevention and treatment of alcohol misuse,[56,57] asking and answering questions about this is often uncomfortable for physicians and patients. For this reason, it is suggested that this question is excluded from the model and the predictAL-9 is used.

The predictAL-10 could be useful in specific contexts, however, such as a longer doctor–patient interview with a climate of mutual trust and empathy or self-administered assessments on a secure website.

Implications for practice

Evidence exists for the effectiveness and cost-effectiveness of screening and brief interventions for hazardous drinkers implemented by primary care professionals (secondary prevention),[7,10] but much less is known about interventions to prevent the onset of hazardous drinking in primary care (primary prevention). The predictAL-10/9, with only 10 or nine items, allows simultaneous screening of current HAD and its prediction at 12 months, which provides an opportunity to carry out both primary and secondary prevention of HAD.

In the present study none of those who developed HAD at 6 months had recovered their status of abstinent or low-risk drinker by 12 months, suggesting that it is important to intervene early through primary prevention. The predictAL-10/9 offers two potential applications:

- a better way of stratifying the at-risk population for inclusion in preventive programmes, and
- the ability to develop personalised preventive programmes based on the overall level of risk and those specific risk factors affecting each person.

The predictAL-10/9 could contribute to the latter just as the predictD risk algorithm is used to prevent the onset of major depression in primary care.[58,59]

REFERENCES

1. Lim SS, Vos T, Flaxman AD, *et al.* A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012; **380(9859):** 2224–2260.
2. World Health Organization, Regional Office for Europe. *Alcohol in the European Union: consumption, harm and policy approaches*. Copenhagen: WHO, 2012.
3. Rehm J, Zatonski W, Taylor B, Anderson P. Epidemiology and alcohol policy in Europe. *Addiction* 2011; **106(Suppl. 1):** 11–19.
4. World Health Organization. *Global status report on alcohol and health — 2014.* Geneva: WHO, 2012.
5. Ministry of Health. *Clinical Commission of the Government Delegation for the National Drug Plan: Report on Alcohol, 2007.* [In Spanish]. Madrid, Centro de Publicaciones del Ministerio de Sanidad y Consumo, 2007.
6. Saitz R. Clinical practice: unhealthy alcohol use. *N Engl J Med* 2005; **352(6):** 596–607.
7. Jonas DE, Garbutt JC, Amick HR, *et al.* Behavioral counseling after screening for alcohol misuse in primary care: a systematic review and meta-analysis for the U.S. Preventive Services Task Force. *Ann Intern Med* 2012; **157(9):** 645–654.
8. Thomas G. *Levels and patterns of alcohol use in Canada. (Alcohol Price Policy Series: Report 1).* Ottawa, ON: Canadian Centre on Substance Abuse, 2012.
9. Drummond C, Oyefeso A, Phillips T, *et al. Alcohol needs assessment research project (ANARP). The 2004 national needs assessment for England.* London: Department of Health and the National Treatment Agency, 2004.
10. Segura-Garcia L, Gual-Solé A, Mestre OM, *et al.* [Detection and handling of alcohol problems in primary care in Catalonia]. [In Spanish]. *Aten Primaria* 2006; **37(9):** 484–490.
11. Watson JM, Crosby H, Dale VM, *et al.* AESOPS: a randomised controlled trial of the clinical effectiveness and cost-effectiveness of opportunistic screening and stepped care interventions for older hazardous alcohol users in primary care. *Health Technol Assess* 2013; **17(25):** 1–158.
12. Muckle W, Muckle J, Welch V, Tugwell P. Managed alcohol as a harm reduction intervention for alcohol addiction in populations at high risk for substance abuse. *Cochrane Database Syst Rev* 2012; **12:** CD006747.
13. Klimas J, Tobin H, Field CA, *et al.* Psychosocial interventions to reduce alcohol consumption in concurrent problem alcohol and illicit drug users. *Cochrane Database Syst Rev* 2014; **12:** CD009269.
14. O'Donnell A, Anderson P, Newbury-Birch D, *et al.* The impact of brief alcohol interventions in primary healthcare: a systematic review of reviews. *Alcohol Alcohol* 2014; **49(1):** 66–78.
15. King M, Marston L, Švab I, *et al.* Development and validation of a risk model for prediction of hazardous alcohol consumption in general practice attendees: the predictAL study. *PLoS One* 2011; **6(8):** e22175.
16. Barrett SP, Tichauer M, Leyton M, Pihl RO. Nicotine increases alcohol self-administration in non-dependent male smokers. *Drug Alcohol Depend* 2006; **81(2):** 197–204.
17. Fenton MC, Geier T, Keyes K, *et al.* Combined role of childhood maltreatment, family history, and gender in the risk for alcohol dependence. *Psychol Med* 2013; **43(5):** 1045–1057.
18. Wellman RJ, Contreras GA, Dugas EN, *et al.* Determinants of sustained binge drinking in young adults. *Alcohol Clin Exp Res* 2014; **38(5):** 1409–1415.
19. Chou KL, Liang K, Mackenzie CS. Binge drinking and Axis I psychiatric disorders in community-dwelling middle-aged and older adults: results from the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). *J Clin Psychiatry* 2011; **72(5):** 640–647.
20. Rubio-Valera M, Pons-Vigués M, Martínez-Andrés M, *et al.* Barriers and facilitators for the implementation of primary prevention and health promotion activities in primary care: a synthesis through meta-ethnography. *PLoS One* 2014; **9(2):** e89554.
21. McBride O, Adamson G, Cheng HG, Slade T. Changes in drinking patterns in the first years after onset: a latent transition analysis of National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) data. *Psychol Addict Behav* 2014; **28(3):** 696–709.
22. Bellón JA, Moreno-Küstner B, Torres-González F, *et al.* Predicting the onset and persistence of episodes of depression in primary care. The predictD-Spain study. methodology. *BMC Public Health* 2008; **8:** 256.
23. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): the TRIPOD Statement. *Ann Intern Med* 2015; **162(1):** 55–63.
24. Babor TF, Higgins-Baddle JC, Saunders JB, Monteiro MG. *The Alcohol Use Disorders Identification Test. guidelines for use in primary care*, 2nd edn. Geneva: World Health Organization, 2001.
25. Rubio Valladolid G, Bermejo Vicedo J, Caballero Sánchez-Serrano MC, *et al.* [Validation of the alcohol use disorders identification test (AUDIT) in primary care]. [In Spanish]. *Rev Clin Esp* 1998; **198(1):** 11–14.
26. Pérula-de Torres LA, Fernández-García JA, Arias-Vega R, *et al.* [Validity of AUDIT test for detection of disorders related with alcohol consumption in women]. [In Spanish]. *Med Clin (Barc)* 2005; **125(19):** 727–730.
27. Pavlou M, Ambler G, Seaman SR, *et al.* How to develop a more accurate risk prediction model when there are few events. *BMJ* 2015; **351:** h3868.
28. Greenland S. Modeling variables selection in epidemiologic analysis. *Am J Public Health* 1989; **79(3):** 340–349.
29. Hernan MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. *Epidemiology* 2004; **15(5):** 615–625.
30. Bellón JA, de Dios Luna J, Moreno-Küstner B, *et al.* Psychosocial and sociodemographic predictors of attrition in a longitudinal study of major depression in primary care: the predictD-Spain study. *J Epidemiol Community Health* 2010; **64(10):** 874–884.
31. Harrell FE. *Regression modelling strategies*. New York, NY: Springer, 2001.
32. Copas JB. Regression, prediction and shrinkage. *J R Statist Soc B* 1983; **45:** 311–354.
33. Cooper H, Hedges LV. *The handbook of research synthesis*. New York, NY: Russell Sage Foundation, 1994.
34. Fluss R, Faraggi D, Reiser B. Estimation of the Youden Index and its associated cutoff point. *Biom J* 2005; **47(4):** 458–472.
35. King M, Walker C, Levy G, *et al.* Development and validation of an international risk prediction algorithm for episodes of major depression in general practice attendees: the PredictD study. *Arch Gen Psychiatry* 2008; **65(12):** 1368–1376.
36. Bellón JA, de Dios Luna J, King M, *et al.* Predicting the onset of major depression in primary care: international validation of a risk prediction algorithm from Spain. *Psychol Med* 2011; **41(10):** 2075–2088.
37. Van Voorhees BW, Paunesku D, Gollan J, *et al.* Predicting future risk of depressive episode in adolescents: the Chicago Adolescent Depression Risk Assessment (CADRA). *Ann Fam Med* 2008; **6(6):** 503–511.
38. Wang J, Sareen J, Patten S, *et al.* A prediction algorithm for first onset of major depression in the general population: development and validation. *J Epidemiol Community Health* 2014; **68(5):** 418–424.
39. Moreno-Peral P, Luna de Dios J, Marston L, *et al.* Predicting the onset of anxiety syndromes at 12 months in primary care attendees. The predictA-Spain study. *PLoS One* 2014; **9(9):** e106370.
40. King M, Bottomley C, Bellón-Saameño JA, *et al.* An international risk prediction algorithm for the onset of generalized anxiety and panic syndromes in general practice attendees: predictA. *Psychol Med* 2011; **41(8):** 1625–1639.
41. Wilson PWF, D'Agostino RB, Levy D, *et al.* Prediction of coronary heart disease using risk factor categories. *Circulation* 1998; **97(18):** 1837–1847.
42. Conroy RM, Pyörälä K, Fitzgerald AP, *et al.* Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur Heart J* 2003; **24(11):** 987–1003.
43. Fuller E, Henderson H, Nass L, *et al. Smoking, drinking and drug use among young people in England in 2012.* London: NatCen Social Research, 2013.
44. Ministry of Health. *Government Delegation for the National Plan on Drugs: survey on drug use in secondary education in Spain 2014/2015.* [In Spanish]. Madrid, Centro de Publicaciones del Ministerio de Sanidad y Consumo, 2015.
45. Lee ML, Yano EM, Wang M, *et al.* What patient population does visit-based sampling in primary care setting represent? *Med Care* 2002; **40(9):** 761–770.
46. Grant BF, Goldstein RB, Chou SP, *et al.* Sociodemographic and psychopathologic predictors of first incidence of DSM-IV substance use, mood and anxiety disorders: results from the Wave 2 National Epidemiologic Survey on Alcohol and Related Conditions. *Mol Psychiatry* 2009; **14(11):** 1051–1066.
47. Galán I, González MJ, Valencia-Martín JL. [Alcohol drinking patterns in Spain: a country in transition]. [In Spanish]. *Rev Esp Salud Publica* 2014; **88(4):** 529–540.
48. Danielsson AK, Wennberg P, Hibell B, Romelsjö A. Alcohol use, heavy episodic drinking and subsequent problems among adolescents in 23 European countries: does the prevention paradox apply? *Addiction* 2012; **107(1):** 71–80.
49. Nazareth I, Walker C, Ridolfi A, *et al.* Heavy episodic drinking in Europe: a cross section study in primary care in six European countries. *Alcohol Alcohol* 2011;

46(5): 600–606.

50. van der Deen FS, Carter KN, McKenzie SK, Blakely T. Do changes in social and economic factors lead to changes in drinking behavior in young adults? Findings from three waves of a population based panel study. *BMC Public Health* 2014; **14:** 928.
51. Harrison ELR, McKee SA. Non-daily smoking predicts hazardous drinking and alcohol use disorders in young adults in a longitudinal U.S. sample. *Drug Alcohol Depend* 2011; **118(1):** 78–82.
52. McKee SA, Weinberger AH. How can we use our knowledge of alcohol-tobacco interactions to reduce alcohol use? *Annu Rev Clin Psychol* 2013; **9:** 649–674.
53. Doyon WM, Thomas AM, Ostroumov A, *et al*. A. Potential substrates for nicotine and alcohol interactions: a focus on the mesocorticolimbic dopamine system. *Biochem Pharmacol* 2013; **86(8):** 1181–1193.
54. Bond JC, Weisner CM, Delucchi K. Alcohol screening and changes in problem drinking behaviors in medical care settings: a longitudinal perspective. *J Stud Alcohol Drugs* 2011; **72(3):** 471–479.
55. Elliott JC, Stohl M, Wall MM, *et al*. The risk for persistent adult alcohol and nicotine dependence: the role of childhood maltreatment. *Addiction* 2014; **109(5):** 842–850.
56. Skinner ML, Kristman-Valente AN, Herrenkohl TI. Adult binge drinking: childhood sexual abuse, gender and the role of adolescent alcohol-related experiences. *Alcohol Alcohol* 2016; **51(2):** 136–141.
57. Sugarman DE, Kaufman JS, Trucco EM, Brown JC, Greenfield SF. Predictors of drinking and functional outcomes for men and women following inpatient alcohol treatment. *Am J Addict* 2014; **23(3):** 226–233.
58. Bellón JA, Conejo-Cerón S, Moreno-Peral P, *et al*. Preventing the onset of major depression based on the level and profile of risk of primary care attendees: protocol of a cluster randomised trial (the predictD-CCRT study). *BMC Psychiatry* 2013; **13:** 171.
59. Bellón JA, Conejo-Cerón S, Moreno-Peral P, *et al*. Intervention to prevent major depression in primary care: a cluster-randomized trial. *Ann Intern Med* 2016; **164(10):** 656–665.
60 Karasek RA, Theorell T. *Healthy work: stress, productivity, and the reconstruction of working life*. New York, NY: Basic Books; 1990.
61. Weich S, Lewis G. Poverty, unemployment, and common mental disorders: population based cohort study. *BMJ* 1998; **317(7151):** 115–119.
62. Jenkinson C, Layte R, Jenkinson D, *et al*. A shorter form health survey: can the SF-12 replicate results from the SF-36 in longitudinal studies? *J Public Health Med* 1997; **19(2):** 179–186.
63. Gandek B, Ware JE, Aaronson NK, *et al*. Cross-validation of item selection and scoring for the SF-12. Health Survey in nine countries: results from the IQOLA Project. International Quality of Life Assessment. *J Clin Epidemiol* 1998; **51(11):** 1171–1178.
64. Spitzer RL, Kroenke K, Williams JBW. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. *JAMA* 1999; **282(18):** 1737–1744.
65. Baca E, Saiz J, Agüera L, *et al*. [Validation of the Spanish version of PRIME-MD: a procedure for diagnosing mental disorders in primary care]. [In Spanish]. *Actas Esp Psiquiatr* 1999; **27(6):** 375–383.
66. Arroll B, Khin N, Kerse N. Screening for depression in primary care with two verbally asked questions: cross sectional study. *BMJ* 2003; **327(7424):** 1144–1146.
67. Robins LN, Wing J, Wittchen HU, *et al*. The Composite International Diagnostic Interview. An epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Arch Gen Psychiatry* 1988; **45(12):** 1069–1077.
68. Rubio-Stipec M, Bravo M, Canino G. [The Composite International Diagnostic Interview (CIDI): an epidemiologic instrument suitable for using in conjunction with different diagnostic systems in different cultures]. [In Spanish]. *Acta Psiquiatr Psicol Am Lat* 1991; **37(3):** 191–204.
69. Reynolds CF, Frank E, Thase ME, *et al*. Assessment of sexual function in depressed, impotent, and healthy men: factor analysis of a Brief Sexual Function Questionnaire for men. *Psychiatry Res* 1988; **24(3):** 231–250.
70. Fink LA, Bernstein D, Handelsman L, *et al*. Initial reliability and validity of the childhood trauma interview: a new multidimensional measure of childhood interpersonal trauma. *Am J Psychiatry* 1995; **152(9):** 1329–1335.
71. King M, Speck P, Thomas A. The Royal Free interview for religious and spiritual beliefs: development and standardization. *Psychol Med* 1995; **25(6):** 1125–1134.
72. Tyrer P. Personality disorder and social functioning. In: Peck DF, Shapiro CM, eds. *Measuring human problems: a practical guide*. Chichester: Wiley & Sons; 1990, 119–142.
73. Qureshi N, Bethea J, Modell B, *et al*. Collecting genetic information in primary care: evaluating a new family history tool. *Fam Pract* 2005; **22(6):** 663–669.
74. Sproston K, Primatesta P. *Health survey for England 2002: a survey carried out on behalf of the Department of Health. Volume 1: the health of children and young people*. London: TSO, 2003.
75. Motrico E, Moreno-Küstner B, Luna JD, *et al*. Psychometric properties of the List of Threatening Experiences–LTE and its association with psychosocial factors and mental disorders according to the different scoring methods. *J Affect Disord* 2013; **150(3):** 931–940.
76. Janssen I, Hanssen M, Bak M, *et al*. Discrimination and delusional ideation. *Br J Psychiatry* 2003; **182:** 71–76.
77. Blaxter M. *Health and lifestyles*. London: Routledge, 1990.
78. Heatherton TF, Kozlowski LT, Frecker RC, *et al*. Measuring the heaviness of smoking: using self-reported time to the first cigarette of the day and number of cigarettes per day. *Br J Addict* 1989; **84(7):** 791–799.
79. Reinert DF, Allen JP. The Alcohol Use Disorders Identification Test: an update of research findings. *Alcohol Clin Exp Res* 2007; **31(2):** 185–199.
80. Gómez A, Conde A, Santana JM, Jorrín A. Diagnostic usefulness of brief versions of Alcohol Use Disorders Identification Test (AUDIT) for detecting hazardous drinkers in primary care settings. *J Stud Alcohol* 2005; **66(2):** 305–308.

Appendix 1. Alcohol abuse and dependence DSM-IV criteria

Alcohol abuse: a maladaptive pattern of alcohol use leading to clinically significant impairment or distress, as manifested by at least one of the following, occurring within a 12-month period:

- recurrent alcohol use resulting in a failure to fulfil major obligations at work, school, or home;
- recurrent alcohol use in situations in which it is physically hazardous (for example, driving); and
- continued alcohol use despite persistent or recurrent social or interpersonal problems caused or exacerbated by the effects of alcohol.

Alcohol dependence: a maladaptive pattern of alcohol use leading to clinically significant impairment or distress, as manifested by at least three of the following at any time in the same 12-month period:

- tolerance;
- withdrawal;
- alcohol is often taken in larger amounts or over a longer period than was intended;
- a persistent desire or unsuccessful efforts to cut down or control alcohol use;
- a great deal of time spent in activities necessary to obtain alcohol, use alcohol, or recover from its effects;
- important social, occupational, or recreational activities are given up or reduced because of alcohol use; and
- use continues despite knowledge of a persistent or recurrent physical or psychological problem that is likely to have been caused or exacerbated by alcohol.

Risk factors for hazardous alcohol drinking

Forty-one potential risk factors were selected, which are described in detail elsewere.[22] A summary of these is given below:

- Sociodemographic factors: (1) age, (2) sex, (3) marital status, (4) occupation, (5) employment status, (6) ethnic group, (7) nationality, (8) country of birth, (9) educational level, (10) income, (11) owner-occupier of their accommodation, (12) living alone or with others.
- Controls, demands, and rewards for (13) paid and (14) unpaid work, using an adapted version of the job content instrument with seven items each.[22,60]
- (15) Debt and financial strain by means of three questions with Likert responses:[61] 1) General financial strain: *'how well would you say you are managing financially these days?'* (4-Likert); 2) Basic financial strain: *'how often does it happen that you do not have enough money to afford the kind of food or clothing you/your family should have?'* (5-Likert); and 3) Coping with debt: *'how much difficulty do you have in meeting the payments of household and other bills?* (6-Likert).
- (16) Physical and (17) mental wellbeing, assessed by the 12-item Short Form (SF-12)[62,63] and (18) a question on the presence of long-standing illness, disability, or infirmity.
- (19) Anxiety disorders using the anxiety section of the Primary Care Evaluation of Mental Disorders (PRIME-MD).[64] The Spanish version of the PRIME-MD can classify patients who test positive for panic attack, generalised anxiety disorder, and other anxiety disorders.[65]
- (20) A screen for lifetime depression based on the first two questions of the Composite International Diagnostic Interview (CIDI).[66]
- (21) Lifetime use of recreational drugs (CIDI).[67,68]
- Brief questions on the quality of (22) sexual and (23) emotional relationships with a partner, adapted from a standardised questionnaire.[69]
- (24) DSM-IV diagnosis of major depression in the preceding 6 months using the CIDI.[67,68]
- (25) A question on taking medication for anxiety, depression, or stress.
- Childhood experiences of (26) physical, (27) emotional, and/or (28) sexual abuse.[70]
- (29) Nature and strength of spiritual beliefs.[71]
- (30) Presence of serious physical or psychological disorder, or substance misuse problems, or any serious disability in persons who were close friends or relations of participants.
- (31) Difficulty getting on with people and maintaining close relationships, assessed using questions from a social functioning scale.[72]
- (32) History of serious psychological problems or (33) suicide in first-degree relatives.[73]
- (34) Satisfaction with the neighbourhood and (35) perceived safety inside/outside the home using questions from the Health Survey for England.[74]
- (36) Threatening events in the preceding 6 months using the List of Threatening Experiences questionnaire.[75]
- (37) Experiences of discrimination in the preceding 6 months on grounds of sex, age, ethnic group, appearance, disability, or sexual orientation, using questions from a European study.[76]
- (38) Adequacy of social support from family and friends.[77]
- (39) Two questions about smoking habits.[78]
- (40) Whether participants had ever had problems with drinking too much alcohol or had ever received treatment for an alcohol problem.
- (41) From the AUDIT,[24] the AUDIT-C was taken out, which contains only three items on alcohol consumption.[79,80]

Management of clustering effect

To test the hierarchical data structure, the likelihood ratio test of the null model was used taking cumulative incidence of hazardous alcohol drinking at 12 months as the dependent variable and health centre as a random factor versus usual logistic regression (χ^2 11.49; P<0.0004). The Intraclass Correlation Coefficient for Health Centre was 0.141 (95% CI = 0.052 to 0.328). The likelihood ratio test of the null model with the variable family physician as a random factor versus usual logistic regression was also significant (χ^2 3.55; P = 0.0298). The Intraclass Correlation Coefficient of the variable family physician was 0.118 (95% CI = 0.036 to 0.327). The likelihood ratio test of the null model was then checked with the health centre and family physician as random factors versus the null model with only health centre (χ^2 0.00; P = 0.9717). It was therefore decided to use multilevel logistic regression with health centre as the random component.

Appendix 2. Model to predict drop-out[a]

Predictors	OR	95% CI	*P*-value
Constant	0.53	0.30 to 0.95	0.032
Province (Granada as reference)			
Zaragoza	1.45	0.89 to 2.34	0.135
Madrid	1.62	0.99 to 2.64	0.053
Logroño (La Rioja)	0.96	0.60 to 1.56	0.876
Majorca	3.09	1.90 to 5.03	<0.001
Las Palmas	2.07	1.16 to 3.70	0.013
Sex (female as reference)			
Male	1.36	1.45 to 1.61	<0.001
Age (range 18–75 years)	0.986	0.978 to 0.994	0.001
Country of birth (Spain as reference)			
Other	1.34	0.96 to 1.87	0.082
Marital status (married as reference)			
Separated	1.08	0.78 to 1.50	0.652
Widowed	1.02	0.75 to 1.37	0.921
Divorced	1.33	0.84 to 2.11	0.223
Single	1.15	0.92 to 1.42	0.221
Employment (employed as reference)			
Unemployed	1.05	0.79 to 1.41	0.729
Retired	1.27	0.98 to 1.65	0.076
Unable to work	0.94	0.70 to 1.27	0.693
Looking after family	0.88	0.70 to 1.11	0.292
Full-time student	0.61	0.37 to 1.03	0.063
Education (beyond secondary as reference)			
Secondary education	1.09	0.84 to 1.42	0.520
Primary education	1.43	1.11 to 1.84	0.005
Incomplete primary education or illiterate	1.81	1.33 to 2.47	<0.001
Housing status (mortgage as reference)			
Owned and paid	0.88	0.74 to 1.05	0.169
Rented	1.57	1.21 to 2.04	0.001
Other	1.08	0.71 to 1.64	0.716
Enough money to afford food or clothing (always as reference)			
Often	1.01	0.83 to 1.24	0.901
Sometimes	0.95	0.74 to 1.20	0.640
Seldom	2.14	1.13 to 4.06	0.020
Never	0.93	0.44 to 1.97	0.854
Satisfaction with the area where you live (very satisfied as reference)			
Satisfied	1.14	0.96 to 1.36	0.142
Neither satisfied nor dissatisfied	1.15	0.91 to 1.48	0.253
Dissatisfied	1.72	1.44 to 2.59	0.009
Very dissatisfied	1.20	0.76 to 1.88	0.441
Cigarette consumption per day (non-smoking as reference)			
<10	1.25	0.98 to 1.60	0.072
10–20	1.10	0.87 to 1.39	0.426
>20	0.97	0.69 to 1.37	0.876
Mental health (SF-12, range 0–100)	0.994	0.989 to 1.001	0.127

[a]*Multilevel logistic regression with health centre and family physician as random components.*

Appendix 3. Weighted and unweighted predictAL-10[a] model by the inverse probability of remaining in the follow-up to 12 months (IPW)

	[b]PredictAL-10 adjusted for IPW			[c]PredictAL-10 not adjusted for IPW		
Risk factors	**OR**	**95% CI**	***P*-value**	**OR**	**95% CI**	***P*-value**
Constant	0.0008	0.0001 to 0.0071	<0.001	0.0011	0.0001 to 0.0093	<0.001
Province						
Granada (Reference)	1.0			1.0		
Zaragoza	2.02	0.49 to 8.37	0.333	2.10	0.67 to 6.58	0.201
Madrid	0.72	0.15 to 3.56	0.690	0.80	0.19 to 3.37	0.764
Logroño (La Rioja)	7.12	2.05 to 24.79	0.002	6.10	2.05 to 18.13	0.001
Majorca	5.32	1.11 to 25.62	0.037	5.12	1.57 to 16.76	0.007
Las Palmas	3.16	0.61 to 16.28	0.170	3.72	0.95 to 14.63	0.060
Sex						
Female (reference)	1.0			1.0		
Male	3.20	1.29 to 7.91	0.012	3.51	2.02 to 6.08	<0.001
Age (range 18–75 years)	0.993	0.972 to 1.015	0.539	0.994	0.963 to 1.025	0.699
AUDIT-C	2.51	1.63 to 3.85	<0.001	2.42	1.59 to 3.71	<0.001
AUDIT-C*Age	0.991	0.984 to 0.999	0.045	0.992	0.983 to 1.001	0.076
Cigarette consumption per day						
Non-smoking (Reference)	1.0					
<10	2.39	1.21 to 4.73	0.012	2.13	1.01 to 4.50	0.046
10–20	1.28	0.51 to 3.18	0.600	1.15	0.51 to 2.60	0.729
>20	3.48	1.31 to 9.27	0.013	3.84	1.74 to 8.50	0.001
Financial strain						
Living comfortably (Reference)	1.0			1.0		
Doing alright	1.94	0.48 to 7.82	0.351	1.54	0.45 to 5.24	0.490
Finding it difficult or very difficult	4.19	0.98 to 17.84	0.053	3.19	0.86 to 11.77	0.082
Ever treated for alcohol problems						
No (Reference)	1.0					
Yes	11.77	1.98 to 70.05	0.007	10.52	2.23 to 49.67	0.003
Sexual abuse in childhood						
No (never) (Reference)	1.0			1.0		
Yes (rarely, sometimes, often, frequently)	5.07	1.71 to 15.09	0.003	4.72	1.65 to 13.45	0.004

[a]Multilevel logistic regression with health centre as a random component. [b]C-Index = 0.886 (95% CI = 0.854 to 0.918). [c]C-index = 0.886 (95% CI = 0.853 to 0.920). Test for the difference: χ^2 (degree of freedom:1) = 0.15; P = 0.691. AUDIT = Alcohol Use Disorders Identification Test. IPW = inverse probability weighting.