TITLE: **Predicting the onset of hazardous alcohol drinking: development and validation of a simple risk algorithm.**

Juan Ángel Bellón[1], Juan de Dios Luna[2], Michael King[3], Irwin Nazareth[4], Emma Motrico[5], María Josefa GildeGómez-Barragán[6], Francisco Torres-González[7], Carmen Montón-Franco[8], Marta Sánchez-Celaya[9], Miguel Ángel Díaz-Barreiros[10], Catalina Vicens[11], and Patricia Moreno-Peral[12]

[1] MD, PhD, Centro de Salud El Palo, Instituto de Investigación Biomédica de Málaga (IBIMA), Departamento de Medicina Preventiva y Psiquiatría, Universidad de Málaga, Spain.

[2] PhD, Departamento de Bioestadística, Universidad de Granada, Spain.

[3] MD, PhD, Division of Psychiatry, Faculty of Brain Sciences, University College London, United Kingdom.

[4] MD, PhD, Department of Primary care and Population Health, University College London, United Kingdom.

[5] PhD, Universidad Loyola Andalucía, Sevilla, Spain.

[6] MD, PhD, Unidad Docente de Medicina Familiar y Comunitaria de La Rioja, Servicio Riojano de la Salud, Logroño, La Rioja, Spain.

[7] PhD, Centro de Investigación Biomédica en Red de Salud Mental, Universidad de Granada, Spain.

[8] MD, PhD, Centro de Salud Casablanca, Instituto Aragonés de Ciencias de la Salud, Zaragoza, Spain. Departamento de Medicina y Psiquiatría, Universidad de Zaragoza, Spain.

[9] MD, PhD, Directora Continuidad Asistencial Hospital Universitario Infanta Sofía, Madrid, Spain.

[10] MD, PhD, Centro de Salud Vecindario, Gerencia de Atención Primaria de Gran Canaria, Servicio Canario de Salud, Las Palmas, Spain.

[11] MD, PhD, Centro de Salud son Serra-La Vileta, Unidad Docente de Medicina Familiar y Comunitaria de Mallorca, Instituto Balear de la Salud, Palma de Mallorca, Illes Balears, Spain.

[12] PhD, Instituto de Investigación Biomédica de Málaga (IBIMA), Spain.

Corresponding author: Juan Ángel Bellón.
Departamento de Medicina Preventiva, Salud Pública y Psiquiatría
Facultad de Medicina, Universidad de Málaga
Campus de Teatinos
29071 Málaga, Spain
Phone: +34 952131515
Fax: +34 952137131
jabellon@uma.es

Abstract

Background: Little is known about the risk of progressing to hazardous alcohol use in people who are 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 6 Spanish provinces, with evaluations at baseline, 6 and 12 months.

Methods: We measured 41 risk factors and used multilevel logistic regression and inverse-probability-weighting to build the risk algorithm. Our outcome was new occurrence of HAD during the 12-month study, as measured by the AUDIT. We recruited 3,954 adult abstinent or low-risk drinkers from the lists of 174 GPs belonging to the 32 health centres.

Results: The "predictAL-10" risk algorithm included just 9 variables (10 questions): province, gender, 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 0.89(95%CI=0.85–0.92). The optimal cut-off had a sensitivity=0.83 and specificity=0.80. The Copas shrinkage factor was 0.96 and calibration plots showed an accurate goodness-of-fit. Excluding childhood sexual abuse from the model (the "predictAL-9"), the c-index was 0.88(95%CI=0.85-0.91), 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 potentially useful for primary prevention of hazardous alcohol drinking.

Keywords: Alcohol use, Clinical Prediction Rule, Primary Health Care

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. To date only one risk algorithm is 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).

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 conditions.[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 men is <170 gr alcohol per week or <110 for women.[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 is that which is already causing damage to health (physical or mental),[7] while alcohol abuse and dependence lead to clinically significant impairment or distress (see Appendix). Around 30% of the population in the US,[6] 24-25% in Canada[8] and the UK[9] or 18% in Spain[10] is 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] We know much less 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 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 (gender, 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] We therefore aimed to develop and internally validate a shorter and simpler risk algorithm to predict the onset of HAD over 12 months in primary care.

METHODS

Design and setting

We undertook a prospective cohort study with evaluations at baseline, 6 and 12 months. Although this cohort was originally recruited with the aim of developing a risk model for the onset of major depression,[22] in this analysis we aimed to predict the onset of HAD.

The method has been described in detail elsewhere[22] and we have followed the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD).[23] The predictAL-Spain study was conducted with the participation of 174 GPs belonging to 32 health centres (mean=5.4; range=1 to 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 4-6 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 two 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 (e.g. psychosis, bipolar), dementia or severe neurological/sensory illness, terminal illness, the person was scheduled to be out of

the city for more than three 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

Our outcome was new occurrence of HAD during the 12-month study. Alcohol use in the preceding six 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 we used an AUDIT cut-off of ≥8 for men and ≥6 for women. Other Spanish researches indicate that a cut-off of ≥8 has a sensitivity and specificity for women and men together of 0.90,[25] and a cut-off of ≥6 for women a sensitivity of 0.90 and a specificity of 0.95.[26] Cronbach's Alpha varies between 0.86[25] and 0.93[26] and test-retest reliability (Intraclass correlation coefficient) is 0.90.[25]

Measurement of potential risk factors

We selected 41 potential risk factors 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 elsewere,[22] and a summary of them is shown in the Appendix.

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 as hazardous alcohol drinkers in our outcome. We performed multilevel logistic regression including health centre as a random component (see Appendix). We estimated the required sample size based on the need for at least 10 outcome events (HAD) per independent variable included in the prediction rule.[27]

We selected variables 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 gender and age were tested. We used inverse probability weighting[29-30] to adjust for a possible attrition bias due to participants lost to follow-up.

We calculated the c-index[31] to estimate the discriminative validity of the final predictAL-10/9 models. To compare the discriminative validity between risk algorithms we performed the test for correlated c-indexes. Prediction models derived with multivariable regression analysis are known to overestimate regression coefficients. We used a calculation proposed by Copas[32] to estimate overfitting of our prediction models. We calculated effect sizes using Hedge's 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, we highlighted the optimal threshold values (cut-off points) where Youden's J statistic (J=Sensitivity+ Specificity-1)[34] was greater. We conducted all analyses using STATA, release 13.1.[35] All reported P values were two-sided.

RESULTS

Of the 6,299 primary care attendees approached, 1,251(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 three months during the 12 months of follow-up. Of the remaining 5,048 patients asked to take part in the study 4,166(82.5%) gave their consent. These were then interviewed at baseline, but 209(5.02%) were hazardous alcohol drinkers (by AUDIT) and 3(0.07%) had a missing diagnosis, so they were also excluded. Thus, our at-risk population comprised 3,954 patients (Figure-1). The patients' socio-demographic characteristics are shown in Table-1. Of the 3,954 patients, 2,667(67.5%) were interviewed at 6 months and 2,301(58.2%) at 12 months. The main baseline variables associated with drop-outs were province (Majorca and Las Palmas), gender (male), lower age, country of birth (outside Spain), lower educational level, never having enough money to afford

food or clothing (basic financial strain), housing status (rented), and dissatisfaction with the area where they lived (see Table-S1 in the appendix). The AUDIT score at baseline was not a predictor of drop-outs.

<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<Table-1 here>>

<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<Figure-1 here>>

Forty-five of those successfully contacted at 6 months had become HAD and a further 31 of those contacted at 12 months had become HAD. All those who were HAD at 6 months were still 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 9 variables (10 questions) (Table-2): province, gender (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=0.854-0.918) and the effect size (Hedges' g) 1.694(95%CI=1.460-1.928). The calibration showed an accurate 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 we removed the potentially sensitive question about childhood sexual abuse from the analysis, we obtained the predictAL-9 model. Compared to 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). However, the c-indexes, Hedges' g and shrinkage factors were similar to the predictAL-10 (Table-3). There was no statistically significant difference between the c-indexes of the predictAL-10 and the predictAL-9 [Chi^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 we entered all the AUDIT questions into the analysis (predictAL-17) but found that this did not improve the c-index over the predictAL-10 [Chi^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 (Table-S2). A calculator of the predictAL-9/10 is available on the "predictplusprevent" web

(http://www.predictplusprevent.com/indexDefi.php?idioma=en). This includes a set of calculators to predict the occurrence of future episodes of depression, anxiety and/or hazardous alcohol drinking in those persons who are not suffering them currently. "Predictplusprevent" also provides information on activities and interventions to prevent them.

<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<Table-2 here>>

<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<Figure-2 here>>

<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<<Table-3 here>>

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.

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 characterize 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 six months). To our 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[36-39] and anxiety syndromes[40-41] in primary care as well as for risk indices for cardiovascular events.[42-43]

Our study has some limitations. Our sample size was not large enough to 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 in other countries are needed. Our study did not include patients below 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 11-15 years, and drinking clearly increases throughout adolescence.[44-45] We had relatively few events (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] However, according to our calibration plot (Figure-2) and the Copas shrinkage factor (0.96) our level of overfitting was minimal. Our study included a large number of GPs and health centres from 6 provinces in southern, central and northern Spain, and only a few patients refused to participate. Therefore our sample may be representative of primary care attendees in Spain, although patients who attend infrequently may have been under-represented.[46] 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, though estimating the incidence was not the aim of the study. Although we had 41.8% dropouts at 12 months, only slight differences were seen between the predictAL-10 models with and without inverse probability weighting (Table-S2), indicating that loss to follow-up was unlikely to lead to attrition bias.[29]

Comparison with existing literature

The 9 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[37] and anxiety syndromes.[40] This was also the case for country in international risk algorithms to predict such disorders,[15,36,41] and suggests that geographical variability must be taken into account in the prediction models. We suggest using the average of the coefficients of the 5 Spanish provinces 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" web.

The results of our study on gender-age and HAD are consistent with the literature. In most countries men tend to drink more than women,[2] and alcohol use occupies third place for men and twelfth for women among risk factors contributing to global disease burden.[1] The incidence of new cases of HAD and alcohol dependence is greater in men between 20-29 years,[47] 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 we also found an interaction 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,[48] while younger people in Europe generally prefer beer, strong spirits and binge drinking.[48-50]

Evidence from longitudinal studies suggests that the perception of difficulty managing changes in living arrangements and individual deprivation are associated with HAD.[51] Daily and non-daily smokers are at a greater risk for hazardous drinking and alcohol use disorders.[16,52] Smoking increases the risk for alcohol misuse and likely has a causal role in this relationship.[53] A biological mechanism underlying the association between alcohol use and smoking has been proposed.[54] Even after being 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.[55]

The three types of child abuse have been associated with alcohol dependence.[17,56] Although childhood sexual abuse could have implications for the course of prevention and treatment of alcohol misuse,[57-58] asking and answering questions about this is often uncomfortable for physicians and patients. For this reason, we have suggested excluding the question from the model and using the predictAL-9. However, the predictAL-10 could be useful in specific contexts, 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 research and 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 9 items, allows us to perform simultaneous screening of current HAD and its prediction at 12 months; which gives us the opportunity to carry out both primary and secondary prevention of HAD. In our study none of those who developed HAD at 6 month 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: 1) a better way of stratifying the at-risk population for inclusion in preventive programs and 2) the ability to develop personalized preventive programs 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.[59-60]

Funding: This study was supported by the Spanish Ministry of Health, The Institute of Health Carlos III, and the European Regional Development Fund (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.

Ethics 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 Malaga, 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.

Competing interest: The authors all declare they have no competing interests.

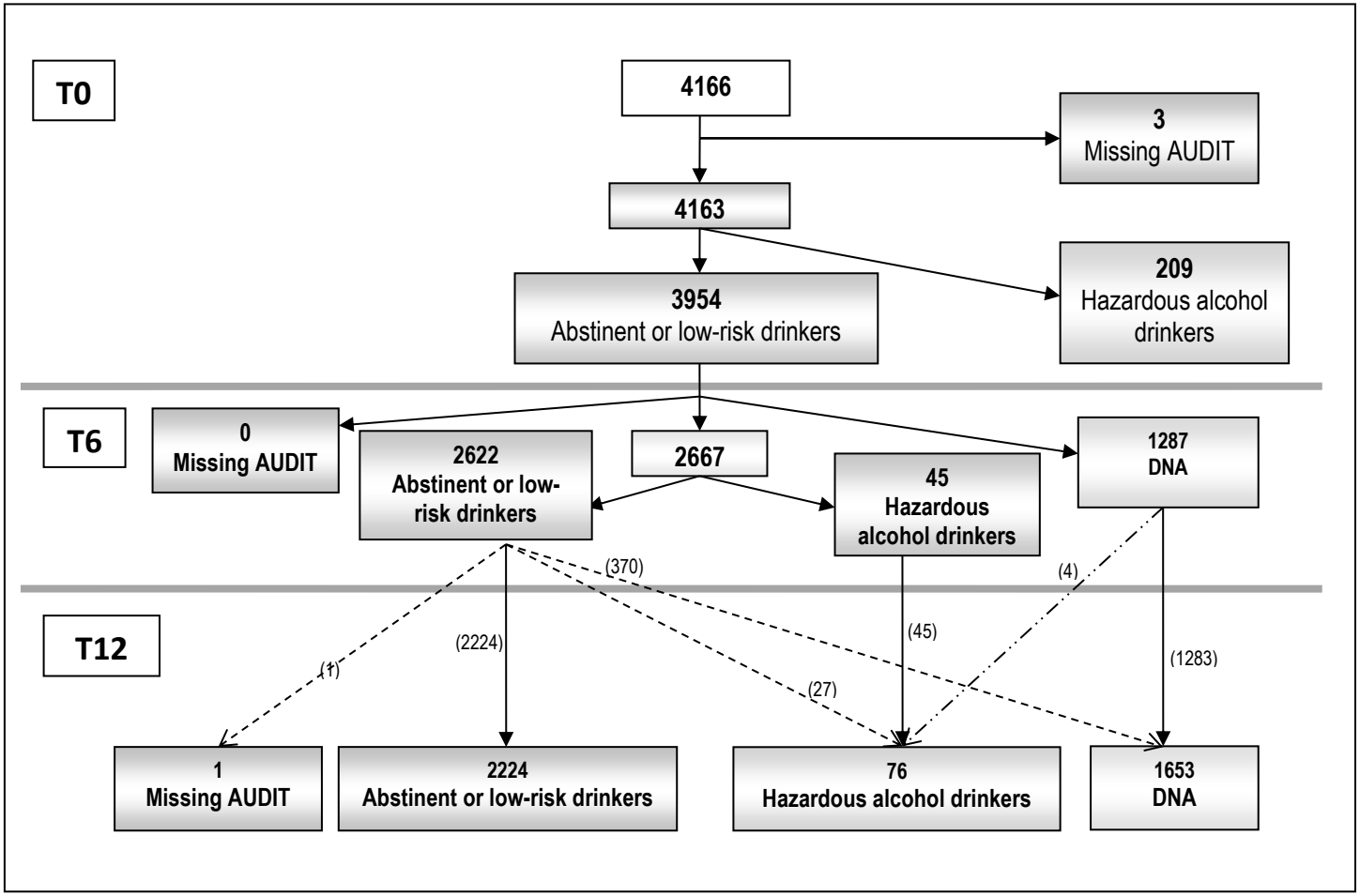
Acknowledgments: The authors thank the Primary Care District of Malaga, particularly Jose Miguel Morales, Javier Navarro and Maximiliano Vilaseca for their support. We also thank the Institute of Health Carlos III (ISCIII), the European Regional Development Fund, the Andalusian Council of Health, the 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 between a number of the authors on this paper. We want especially to thank all the interviewers, patients and Primary Care Physicians for their participation in the predictAL-10 Study.

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:** 2224-60.
2. World Health Organization, Regional Office for Europe. Alcohol in the European Union: consumption, harm and policy approaches. Copenhagen; 2012.
3. Rehm J, Zatonski W, Taylor B. Epidemiology and alcohol policy in Europe. *Addiction* 2011; **106(Suppl. 1):** 11-9.
4. World Health Organization. Global status report on alcohol and health – 2014 ed.
5. Ministry of Health. Clinical Commission of the Government Delegation for the National Drug Plan: Report on Alcohol, 2007. [Report in Spanish]
6. Saitz R. Clinical practice: Unhealthy alcohol use. *N Engl J Med* 2005; **352:** 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:** 645-54.
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. Contract No.: Document Number.
10. Segura-Garcia L, Gual-Solé A, Mestre OM, Bueno-Belmonte A, Colom-Farrana J. Detection and handling of alcohol problems in primary care in Catalonia [Article in Spanish]. *Aten Primaria* 2006; **37:** 484-90.
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, Svab 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:** e22175.
16. Barrett SP, Tichauer M, Leyton M, Pihl RO. Nicotine increases alcohol self-administration in non-dependent male smokers. *Drug and Alcohol Dependence* 2006; **81:** 197-204.
17. Fenton MC, Geier T, Keyes K, Skodol AE, Grant BF, Hasin DS. Combined role of childhood maltreatment, family history, and gender in the risk for alcohol dependence. *Psychol Med* 2013; **43:** 1045-57.
18. Wellman RJ, Contreras GA, Dugas EN, O'Loughlin EK, O'Loughlin JL. Determinants of sustained binge drinking in young adults. *Alcohol Clin Exp Res* 2014; **38:** 1409-15.
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:** 640-7.
20. Rubio-Valera M, Pons-Vigués M, Martínez-Andrés M, Moreno-Peral P, Berenguera A, Fernández A. 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:** 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:** 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 PREDICT-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. World Health Organization. The Alcohol Use Disorders Identification Test Guidelines for Use in Primary Care, 2nd Edition, 2001.
25. Rubio Valladolid G, Bermejo Vicedo J, Caballero Sánchez-Serrano MC, Santo-Domingo Carrasco J. Validation of the alcohol use disorders identification test (AUDIT) in primary care. *Rev Clin Esp* 1998; **198:** 11-14.
26. Pérula-de Torres LA, Fernández-García JA, Arias-Vega R, Muriel-Palomino M, Márquez-Rebollo E, Ruiz-Moral R. Validity of AUDIT test for detection of disorders related with alcohol consumption in women. *Med Clin (Barc)* 2005; **125:** 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:** 340-9.
29. Hernan MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. *Epidemiology* 2004; **15:** 615-25.
30. Bellón JA, Luna JD, 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:** 874-84.
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:** 311-54.

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:** 458-72.
35. StataCorp. Stata statistical software: Release 13. In Station C, editor. Texas: StataCorp LP; 2013.
36. 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:** 1368-76.
37. Bellón JA, Luna JD, 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:** 2075-88.
38. Van Voorhees BW, Paunesku D, Gollan J, Kuwabara S, Reinecke M, Basu A. Predicting future risk of depressive episode in adolescents: the Chicago Adolescent Depression Risk Assessment (CADRA). *Ann Fam Med* 2008; **6:** 503-11.
39. Wang J, Sareen J, Patten S, Bolton J, Schmitz N, Birney A. A prediction algorithm for first onset of major depression in the general population: development and validation. *J Epidemiol Community Health* 2014; **68:** 418-24.
40. Moreno-Peral P, Luna JD, 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:** e106370.
41. 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:** 1625-39.
42. Wilson PWF, D'Agostino RB, Levy D, Belanger AM, Silvershatz H, Kannel B. Prediction of coronary heart disease using risk factor categories. *Circulation* 1998; **97:** 1837-47.
43. Conroy RM, Pyorala K, Fitzgerald AP, *et al*. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur Heart J* 2003; **24:** 987-1003.
44. Fuller E, Henderson H, Nass L, Payne C, Phelps A, Ryley A. Smoking, Drinking and Drug Use Among Young People in England in 2012. London: NatCen Social Research; 2013.
45. Ministry of Health. Government Delegation for the National Plan on Drugs: Survey on drug use in secondary education in Spain 2014/2015. [Report in Spanish]
46. Lee ML, Yano EM, Wang M, Simon BF, Rubenstein LV. What patient population does visit-based sampling in primary care setting represent? *Med Care* 2002; **40:** 761-70.
47. 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:** 1051-66.
48. Galán I, González MJ, Valencia-Martín JL. Alcohol drinking patterns in Spain: a country in transition. *Rev Esp Salud Publica* 2014; **88:** 529-40.
49. 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:** 71-80.
50. 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:** 600-6.
51. 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.
52. 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:** 78–82.
53. McKee SA, Weinberger AH. How can we use our knowledge of alcohol-tobacco interactions to reduce alcohol use? *Annu Rev Clin Psychol* 2013; **9:** 649-74.
54. Doyon WM, Thomas AM, Ostroumov A, Dong Y, Dani J. A. Potential substrates for nicotine and alcohol interactions: a focus on the mesocorticolimbic dopamine system. *Biochem Pharmacol* 2013; **86:** 1181-93.
55. 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**: 471-9.
56. 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:** 842-50.
57. 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:** 136-41.
58. 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:** 226-33.
59. 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.
60. 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:** 656-65.

Figure 1: Flow chart of patients through the predictAL-10/9 study and numbers becoming hazardous alcohol drinkers.



AUDIT (Alcohol Use Disorders Identification Test). Hazardous alcohol drinking [men: AUDIT ≥ 8; women: AUDIT ≥ 6]. Abstinent or low-risk drinkers [men: AUDIT <8; women: AUDIT <6]; DNA, did not attend; T0: at baseline, T6: at 6 months, and T12: at 12 months.

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

Demographic characteristics	Granada	Saragossa	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 (standard deviation)	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)
Gender, 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 (5.90)	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)
Ethnicity, 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 ethnicity	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 alright	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)

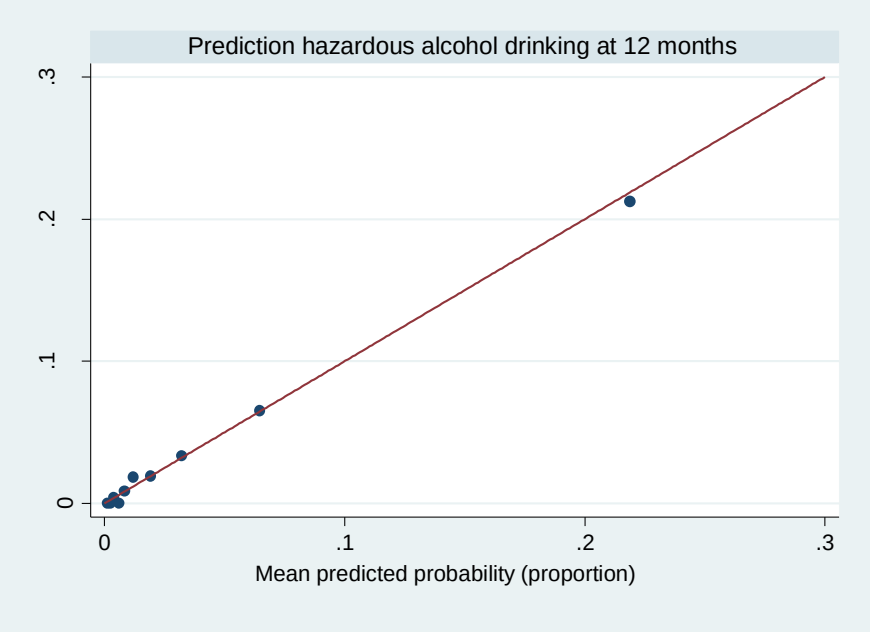
Table 2: The predictD-AL-10 and predictAL-9 models(¶) to predict the onset of hazardous alcohol drinking at 12 months.

Risk factors	Incidence of hazardous alcohol drinking		PredictAL-10 [a] (N=2264)				PredictAL-9 [b] (N=2278) (excluding childhood sexual abuse from the model)			
	No	Yes (%)	OR	OR [c]	95% CI	p	OR	OR [d]	95% CI	p
Constant			0.0008	0.0009	0.0001 - 0.0071	<0.001	0.0011	0.0012	0.0001 - 0.0079	<0.001
Province										
Granada (Reference)	499	7 (1.38)	1.0	1.0			1.0	1.0		
Saragossa	447	16 (3.46)	2.02	1.94	0.49 – 8.37	0.333	1.80	1.73	0.49 – 6.70	0.379
Madrid	406	4 (0.98)	0.72	0.69	0.15 – 3.56	0.690	0.66	0.64	0.15 – 2.93	0.585
Logroño (La Rioja)	429	29 (6.33)	7.12	6.84	2.05 – 24.79	0.002	6.46	6.22	2.11 – 19.79	0.001
Majorca	258	13 (4.80)	5.32	5.11	1.11 – 25.62	0.037	4.74	4.57	1.08 – 20.84	0.040
Las Palmas	185	7 (3.65)	3.16	3.03	0.61 – 16.28	0.170	3.09	2.96	0.62 – 15.41	0.168
Gender										
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 – 7.91	0.012	2.82	2.72	1.62 – 6.83	0.022
Age (range 18-75 years)			0.993	0.953	0.972 – 1.015	0.539	0.989	0.952	0.969 – 1.010	0.316
AUDIT-C			2.51	2.41	1.63 – 3.85	<0.001	2.37	2.28	1.57 – 3.59	<0.001
AUDIT-C*Age (†)			0.991	0.951	0.984 – 0.999	0.045	0.993	0.956	0.985 – 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 – 4.73	0.012	2.42	2.33	1.20 – 4.88	0.014
10-20	200	11 (5.21)	1.28	1.23	0.51 – 3.18	0.600	1.38	1.33	0.57 – 3.31	0.475
>20	87	12 (12.1)	3.48	3.34	1.31 – 9.27	0.013	3.59	3.46	1.33 – 9.68	0.012
Financial strain										
Living comfortably (Reference)	183	3 (1.61)	1.0	1.0			1.0	1.0		
Doing alright	1634	56 (3.31)	1.94	1.86	0.48 – 7.82	0.351	2.07	1.99	0.50 – 8.51	0.313
Finding it difficult or very difficult	405	17 (4.03)	4.19	4.02	0.98 – 17.84	0.053	4.66	4.49	1.08 – 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 – 70.05	0.007	13.19	12.70	2.61 – 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 – 15.09	0.003				

(¶) Multi-level logistic regression with health center as a random component and weighting for the inverse probability of remaining in the follow-up to 12 months. (†) Likelihood-ratio test for the interaction: Chi2 (degree of freedom:1)=5.84;p=0.0157. [a] Discriminative validity: c-index: 0.886 (0.854 - 0.918) and effect size (Hedges' g): 1.694 (95%CI:1.460 - 1.928). [c] Overfitting estimate: Copas' shrinkage factor = 0.960. [b] Discriminative validity: c-index: 0.880 (95%CI:0.847 - 0.913) and effect size (Hedges' g): 1.658 (95%CI:1.425 - 1.892). [d] Overfitting estimate: Copas' shrinkage factor = 0.963.[c]

Figure 2: Calibration plots (mean predicted probability against observed probability of hazardous alcohol drinking within deciles of predicted risk) of the predictAL10/9 risk algorithms.

The predictAL-10



The predictAL-9

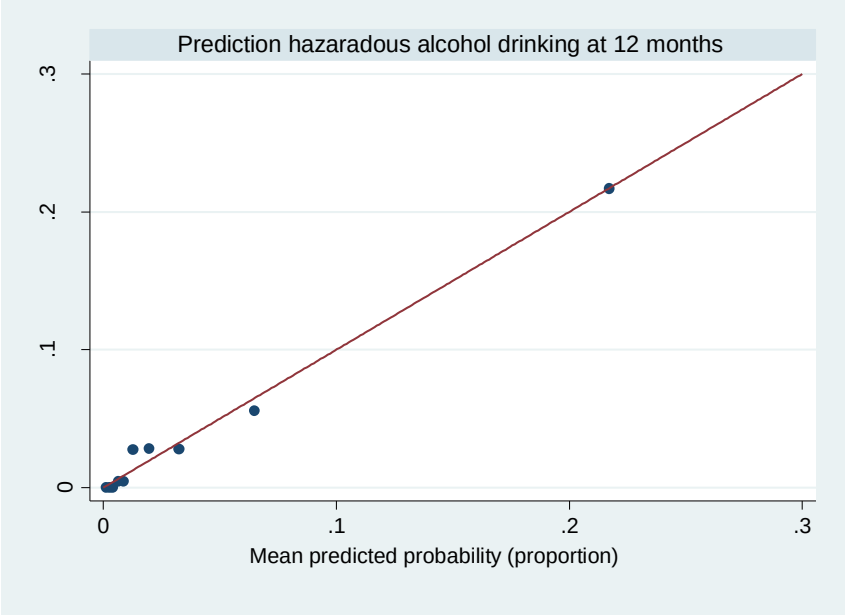


Table 3: Discriminative validity of the clinical rule predictions for the onset of hazardous alcohol drinking over 12 months in primary care.

Risk algorithms	Number of items	N	C-index (95%C.I.)	Hedges's g (95%C.I.)	Predicted probability*	Frequency** N (%)	Sensitivity Specificity	LR+ LR-	PPV NPV	Shrinkage factor ¶
PredictAL-10[a]	**10**	**2,264**	**0.886 (0.854 - 0.918)**	**1.694 (1.460 - 1.928)**	**≥ 2.72%**	**495 (21.86)**	**0.83 0.80**	**4.15 0.21**	**0.13 0.99**	**0.9595**
PredictAL-17 [b]	17	2,264	0.886 (0.853 - 0.919)	1.729 (1.495 - 1.963)	≥ 2.92%	454 (20.05)	0.80 0.82	4.44 0.25	0.13 0.99	0.9595
PredictAL-7[c]	7	2,264	0.819 (0.772 - 0.866)	1.292 (1.052 - 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 [a]	**9**	**2,278**	**0.880 (0.847 - 0.913)**	**1.658 (1.425 - 1.892)**	**≥ 3.01%**	**475 (20.85)**	**0.79 0.81**	**4.16 0.26**	**0.13 0.99**	**0.9629**
PredictAL-16 [b]	16	2,278	0.883 (0.850 - 0.916)	1.697 (1.463 - 1.931)	≥ 2.78%	483 (21.20)	0.80 0.81	4.21 0.25	0.13 0.99	0.9639
PredictAL-6 [c]	6	2,278	0.803 (0.752 - 0.854)	1.273 (1.042 - 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	2,288	0.775 (0.721 - 0.830)	1.211 (0.980 - 1.442)	≥ 2.34%	517 (22.60)	0.75 0.68	4.16 0.26	0.07 0.89	0.9819
AUDIT	10	2,288	0.781 (0.725 - 0.836)	1.254 (1.042 - 1.485)	≥ 2.20%	525 (22.95)	0.75 0.68	4.21 0.25	0.07 0.89	0.9822

***C.I.**: Confidence Interval; **LR+**: positive likelihood ratio; **LR-**: negative likelihood ratio; **PPV**: positive predictive value; **NPV**: negative predictive value;*
¶ Copas shrinkage factor estimates overfitting of the prediction models (shrinkage = 1 indicates that there is no overestimation).
[a] Including the AUDIT-C (3 items); [b] including the AUDIT (10 items); [c] excluding any AUDIT.
** Predicted probability of hazardous alcohol drinking at 12 months, cutoff point where Youden's J statistic (J = Sensitivity + Specificity -1) was greater: "optimal threshold".*
*** Number of primary care attendees above the optimal threshold.*

PRINT VERSION

Background

Alcohol use occupies fifth place among risk factors contributing to worldwide global disease burden.[1] Many risk factors are associated with the onset of hazardous or harmful alcohol drinking,[14-17] but so far only one risk algorithm taking into account their combined effect has been published; the predictAL,[13] which requires the patient having to answer 29 questions to calculate their risk at 6 months. We aimed 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.

Method

The predictAL-Spain is a prospective cohort study which was conducted with the participation of 174 GPs belonging to 32 health centres distributed throughout Spain (in six provinces), with evaluations at baseline, 6 and 12 months. Random samples of attendees from the appointment lists of the 174 GPs were taken between October 2005 and February 2006. Exclusion criteria were an inability to understand or speak Spanish, severe mental disorder (e.g. psychosis, bipolar), dementia or severe neurological/sensory illness, terminal illness, the person was scheduled to be out of the city for more than three 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. We measured 41 risk factors[22] and used multilevel logistic regression and inverse-probability-weighting[29-30] to build the risk algorithm. We used a calculation proposed by Copas[32] to estimate overfitting of our prediction models and performed calibration plots. Our outcome was new occurrence of HAD during the 12-month study, as measured by the AUDIT.[24-26]

Results

Of the 6,299 primary care attendees approached, 1,251(19.9%) were excluded for having at least one exclusion criteria. Of the remaining 5,048 patients asked to take part in the study 4,166(82.5%) gave their consent. These were then interviewed at baseline, but 209(5.02%) were hazardous alcohol drinkers (by AUDIT) and 3(0.07%) had a missing diagnosis, so they were also excluded. Thus, our at-risk population comprised 3,954 patients, of which 2,667(67.5%) were interviewed at 6 months and 2,301(58.2%) at 12 months. Forty-five of those successfully contacted at 6 months had developed HAD and a further 31 of those contacted at 12 months had become HAD. All those developing HAD by 6 months were still HAD by 12.

The "predictAL-10" risk algorithm included just 9 variables (10 questions): province, gender, 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 (see **table**). The c-index was 0.89(95%CI=0.85–0.92). The optimal cut-off had a sensitivity=0.83 and specificity=0.80. The Copas shrinkage factor was 0.96 and calibration plots showed an accurate goodness-of-fit. Excluding childhood sexual abuse from the model (the "predictAL-9"), the c-index was 0.88(95%CI=0.85-0.91), sensitivity=0.79 and specificity=0.81. There was no statistically significant difference between the c-indexes of predictAL-10 and predictAL-9.

Discussion

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.

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 characterize 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 six months). To our 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[36-39] and anxiety syndromes[40-41] in primary care as well as for risk indices for cardiovascular events.[42-43]

Our sample size was not large enough to address external validation in this study, and new validations of the predictAL-9/10 in other countries are needed. Although we had 41.8% dropouts at 12 months, only slight differences were seen between the predictAL-10 models with and without inverse probability weighting, indicating that loss to follow-up is unlikely to lead to attrition bias.[29]

The predictAL-10/9, with only 10 or 9 items, allows us to perform simultaneous screening of current HAD and its prediction at 12 months; which gives us the opportunity to carry out both primary and secondary prevention of HAD. In our study none of those who developed HAD at 6 month had recovered their status of 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: 1) a better way of stratifying the at-risk population for inclusion in preventive programs and 2) the ability to develop personalized preventive programs based on the overall level of risk and those specific risk factors that affect each person. The predictAL-10/9 could contribute to the second of these applications just as the predictD risk algorithm is used to prevent the onset of major depression in primary care.[59-60]

Table: The predictD-AL-10 and predictAL-9 models(¶) to predict the onset of hazardous alcohol drinking at 12 months.

Risk factors	PredictAL-10 [a] (N=2264)		PredictAL-9 [b] (N=2278)	
	OR	95% CI	OR	95% CI
Constant	0.0008**	0.0001 - 0.0071	0.0011**	0.0001 - 0.0079
Province				
Granada (Reference)	1.0		1.0	
Saragossa	2.02	0.49 – 8.37	1.80	0.49 – 6.70
Madrid	0.72	0.15 – 3.56	0.66	0.15 – 2.93
Logroño (La Rioja)	7.12**	2.05 – 24.79	6.46**	2.11 – 19.79
Majorca	5.32*	1.11 – 25.62	4.74*	1.08 – 20.84
Las Palmas	3.16	0.61 – 16.28	3.09	0.62 – 15.41
Gender				
Female (Reference)	1.0		1.0	
Male	3.20*	1.29 – 7.91	2.82*	1.62 – 6.83
Age (range 18-75 years)	0.993	0.972 – 1.015	0.989	0.969 – 1.010
AUDIT-C	2.51**	1.63 – 3.85	2.37*	1.57 – 3.59
AUDIT-C*Age (†)	0.991*	0.984 – 0.999	0.993	0.985 – 1.000
Cigarette consumption per day				
Non-smoking (Reference)	1.0		1.0	
<10	2.39*	1.21 – 4.73	2.42*	1.20 – 4.88
10-20	1.28	0.51 – 3.18	1.38	0.57 – 3.31
>20	3.48*	1.31 – 9.27	3.59*	1.33 – 9.68
Financial strain				
Living comfortably (Reference)	1.0		1.0	
Doing alright	1.94	0.48 – 7.82	2.07	0.50 – 8.51
Finding it difficult or very difficult	4.19	0.98 – 17.84	4.66*	1.08 – 20.03
Ever treated for alcohol problems				
No (Reference)	1.0		1.0	
Yes	11.77**	1.98 – 70.05	13.19**	2.61 – 66.63
Childhood sexual abuse				
No (never) (Reference)	1.0			
Yes (rarely, sometimes, often, frequently)	5.07**	1.71 – 15.09		

* p<0.05; ** p<0.01.

(¶) Multi-level logistic regression with health centre as a random component and weighting for the inverse probability of remaining in the follow-up to 12 months. (†) Likelihood-ratio test for the interaction: Chi2 (degree of freedom:1)=5.84;p=0.0157. [a] Discriminative validity: c-index: 0.886 (0.854 - 0.918) and effect size (Hedges' g): 1.694 (95%CI:1.460 - 1.928). [b] Discriminative validity: c-index: 0.880 (95%CI:0.847 - 0.913) and effect size (Hedges' g): 1.658 (95%CI:1.425 - 1.892).